

Europe's embarrassing growth

The European Communities seem at a loss to know how to respond to new membership applications. Some (like that of Sweden) seem too much of the same, others (mostly from Eastern Europe) raise economic problems Europe is unwilling to face.

THE European Communities (EC) are busily preparing for the arrival of their single market at the beginning of 1993 — there is, for example, a scheme to build a whole chain of bonfires at the literal beginning of that year. Before then, the EC may have won some self-respect as a cryptopower on the international scene from its efforts to prevent fighting between Serbs and Croats in Yugoslavia. Yet the EC is still in a thorough muddle about its own more distant future. It has seemed from Brussels already bad enough that after the accession of four countries (Britain, Spain, Portugal and Greece) in just 20 years, the queue of would-be members is still at least three (Turkey, Sweden and Austria), perhaps as many as five (with Norway and possibly Switzerland), and is likely to be swollen more rapidly than it can be made to shrink by clamourings from the East. What is to be done?

The underlying dilemma is not new. It is also a difficulty that can afflict any kind of club. Those who belong already tend to like things the way they are, and prefer not to see them changed. The newest members are often the most zealous in their defence of the status quo. But those who do not yet belong feel neglected, even scorned, and try to insist that they have a right of membership. In the EC, those who wish to contain the membership (at least for the time being) argue that the first task is to deepen the relationship between the EC's members — a process certainly likely to be impeded by growth. But others hold that, just as one of the motives for the foundation of the EC was the need to bring previously warring states within a common political framework, so Europe now has an obligation (which is also an opportunity) to extend that framework to the East. The difference between the 'deepeners' and the 'broadeners' is longstanding and unlikely to be quickly or permanently resolved. That is why the EC should be looking for ways of blunting the dilemma.

The place to start is with an understanding of what the would-be members want. Their motives are not identical. Sweden and Austria, for example, see the issue as one of simple economics. Much of their external trade is with countries elsewhere in Europe. To be left out might be a hazard (if, for example, the EC turns out to be a high-tariff region), while the price of membership is not high (especially because Ireland serves as a precedent of neutrality). Others — Turkey as well as the states of Eastern Europe — on the other hand would expect from the conditions of

past membership that they would enjoy the benefits of the Common Agricultural Policy and, eventually, capital funds for development.

This may be mistaken. The states of Central Europe are evidently in a separate category. The last thing they want at this stage is to be faced with the full rigours of the competition by which the EC hopes to make itself prosperous over the decades ahead. Instead, they need markets for simple manufactures and for the food that they produce. Unless these opportunities are offered, they may have no choice but to slink back into the club from which they have only recently emerged. Therefore, the EC should say now that there will be membership at some point in the future if they still want it.

Those who are fond of talking of "Europe from the Atlantic to the Urals" should appreciate that there is no real sense in any other policy, however much discontent it might engender among manufacturers in the West. The European Development Bank, which came into being earlier this year (and which is a good idea), intends to use capital from the West to revitalize industry in the East. But what use will that be if the East has no customers for the products its more efficient industry will be turning out? Agriculture is a greater mess. By the end of this year, the European Common Agricultural Policy will have to be amended to satisfy hostile negotiators in the General Agreement on Tariffs and Trade, but it is not yet sufficiently clear in Brussels that if the Common Agricultural Policy were simply scrapped, the immediate difficulties of Central European states would vanish as if by magic. Is not that a sufficient incentive to take a step long overdue for its own sake and in the interests of seemly logic?

Meanwhile, Europe has the cultural card to play. If Europe is capable of becoming the integral entity of which the enthusiasts speak, should not Brussels take the lead in bringing some of that about? This is not a plea that there should be more itinerant groups of folk dancers wearing national dress, but that Brussels should deliberately set out to ensure that ideas arising anywhere in the academic or research communities of Europe are free to travel — and with their originators.

It is slightly shameful that so far most of the running in the cultural field has been made by commercial organizations, publishers and television companies. Yet not much money would be needed — merely a little institution building and the right kinds of speeches. □



The transmission of AIDS

Data that mothers can transmit HIV to infants through breast feeding add disturbing new insight into the mechanism of transmission.

In Kigali, Rwanda, physicians have been following the development of infants born to 212 mothers who were seronegative for HIV (human immunodeficiency virus) at the time of delivery. "After a mean follow-up of 16.6 months, 16 of the 212 mothers became seropositive for HIV-1. Of their 16 infants, 9 became seropositive," according to a paper in *The New England Journal of Medicine* (325, 593-598; 29 August 1991). Each of the infants was breast fed; mothers' milk was the route of HIV transmission.

These data add one more discouraging piece of evidence to the hypothesis that AIDS is more easily transmissible than people would like to believe. Although the idea that HIV might be transmitted through breast milk is not novel, the possibility of it ever happening has consistently been described as "rare" by AIDS researchers and epidemiologists. However, the seroconversion rate in Rwanda alone is in the range of 4.7 to 6.3 per cent a year, putting many children at risk. In the United States and Europe, where the virus is still not widespread among women, the data from Africa may not seem threatening. However, speculation that HIV may enter the heterosexual population (as it already has in Africa and Asia) needs to be taken seriously.

Thus it is crucial to know more about the mechanism of HIV transmission. At the seventh international conference on AIDS, held in Florence in June, new attention was paid to preliminary data showing that HIV may penetrate dendritic cells in genital (and possibly oral) mucosa. If this is true, then the belief that HIV transmission requires blood-to-blood contact falls on its face.

For the present, there is no suggestion that poor women in Rwanda or elsewhere should give up breast feeding, but that is mainly because they lack safe alternatives for feeding their infants. The dangers of non-sterile formula outweigh the statistical probability of contracting AIDS. But that is not the point.

The new data about mothers' milk merely adds to the urgency of learning more about mucosal transmission because the most likely explanation for HIV transmission through breast feeding is that the virus penetrates the mucosal lining of the gastrointestinal tract of infants. It may be necessary to abandon the comforting notion that HIV is extraordinarily difficult to transmit.

The unanswered questions about mucosal transmission are pertinent to a heated debate in the United States about the duties of HIV-infected medical professionals — particularly dentists and surgeons — who come into contact with patients' bodily fluids. The US government has asked professional groups to develop lists of procedures that HIV carriers or people with full-blown AIDS should

avoid in medical or dental practice. It is a sensible, if conservative, approach to the unthinkable — contracting AIDS from one's doctor.

However, nearly 40 groups of health care professionals have declined to draw up guidelines on the grounds that the likelihood of doctor-patient transmission is so remote that guidelines are not necessary. It is true that the case of a Florida dentist who transmitted AIDS to a now-dying patient is the only known instance of such transmission. But it is also a matter of record that more than 6,000 health care workers are now known to be HIV-positive and the US Centers for Disease Control say that number probably is low.

The fact is that at present we do not know the full extent of the risk. But emerging data suggest that the recalcitrant groups of health professionals should think again. □

More genome ethics

Now Baroness Warnock has joined the chorus saying that even knowledge can be ethically dangerous.

THE British Association for the Advancement of Science seems to have had a traditionally routine annual meeting at Plymouth last week. Industrialists complained that other industrialists were not spending enough on research and development, while environmentalists complained about all others but themselves. But something must be done about the way that meetings such as this are increasingly used to amplify the over-simple message that the human genome projects abound with ethical difficulties. At Plymouth, Baroness Warnock, the chairman of the committee whose largely sensible report became the basis for British legislation on human embryo research, was asking that there should now be legislation to restrict access to human genome data by insurance companies and employers.

What exactly are people afraid of? Insurance has been much discussed. Consider employment, and suppose that genome research some day describes an allelic version of a normal human gene that makes its carriers susceptible to the carcinogenic effect of a chemical used widely in some industrial process. What should happen then? Should plants using the process be shut down on the grounds that susceptible people might be employed in them? Or should employers be encouraged to discriminate among their potential employees, not hiring those discovered to be susceptible?

We do not have to wait for the outcome of the genome projects to know what will happen. There is an inherited susceptibility to vinyl chloride monomer, but the attempts by employers in the United States to avoid exposing susceptible people to danger have been met with a flurry of lawsuits. Does that imply that the right to work in the manufacture of PVC is one of the inalienable rights of man? Would it not be profitable to keep the plants going and to use part of the economic wealth created to compensate those with the bad luck to be susceptible? □

Japanese science agency targets space, genome

■ Module for US space station is a centrepiece

■ 'Inspiration' for interdisciplinary research

Tokyo

JAPANESE space efforts are set to continue their rapid growth, as the Science and Technology Agency (STA) has requested a 10 per cent increase in funds for that programme over the next fiscal year. The agency also says it is placing new emphasis on research on the human genome with a proposed jump in funding of more than 50 per cent, although the total investment remains small by international standards.

Those are the highlights of the budget request submitted by the agency last week to the Ministry of Finance. The request will be trimmed in negotiations with the finance ministry before it is submitted to the Diet early next year, but no major changes are likely.

The budget is dominated by space and nuclear power, which account for more than 80 per cent of the proposed expenditure. The large increase requested for space will only serve to strengthen the view, held by other government organizations, that the agency is primarily an organization for the development of space and nuclear power, rather than science and technology in general as its name implies.

Compared with other government ministries and agencies, which must operate under very strict budget limits, the STA has greater freedom to expand its budget — hence the comparatively large overall increase of 6.4 per cent requested. In particular, the agency has flexible funding limits on international projects run under intergovernmental agreements, the prime example of which is the US space station. Japan's commitment to it adds enormously each year to the agency's outlay for space, leading even some agency officials to say that the space budget is becoming lopsided.

Most of the more than ¥13,000 million (\$100 million) increase in expenditure for space goes to development of the Japanese module for the US space station, for which the agency has requested ¥28,236 million, up from ¥17,958 million this year. There is also a large increase for the Advanced Earth Observation Satellite, with a proposed budget of ¥11,833 million, or almost double that of this year. The increase for the Earth-observing satellite, which will carry sensors built by the United States and France as well as Japan, accounts for a large proportion of the increase in the agency's proposed outlay for the 'Green Planet Project', which aims at observing the global environment.

The budget request for human genome research has been increased to ¥1,520 million (\$11 million). About ¥280 million (\$2 million) in new funds will be allocated to establishing an international computer link between the Japan Information Center of Science and Technology and the genome data base at Johns Hopkins University in the United States. The Institute of Physical and Chemical Research will be given a similar boost in funds to ¥861 million (\$6.4 million) to expand preliminary research on the human genome.

The genome research will focus on mapping and the build-up of DNA libraries, and the development of information-processing systems and DNA sequencing technology, before embarking on a full-blown project to sequence the human genome from 1995, agency officials say. But ambitious plans by some agency members to establish a new semi-private organization to carry out the work have been postponed until nearer the time of the second phase.

Among the STA's new projects is a

small initiative called 'inspiration', which is allocated ¥285 million (\$2.1 million) in its first year to support interdisciplinary research in the private sector, government and the universities. The funds will be used for symposia, conferences, the exchange of researchers between organizations both within and outside Japan, and for the planning of a new 'international plaza' in Tsukuba science city. Agency officials hope to increase the budget for inspiration considerably in coming years.

The new Sakigake project,

which is intended to provide three-year grants of about ¥20 million (\$150,000) a year to young researchers from universities, government and industry, will be given a large increase to ¥1,175 million. This will allow the agency to award 24 new grants next year in addition to 36 soon to be awarded this year.

The agency has been swamped with more than 460 Sakigake grant applications in the past few weeks, about half of them from universities and university-related institutes. But it remains to be seen if the Ministry of Education, Science and Culture, a bitter rival of the Science and Technology Agency, will allow the university researchers under its jurisdiction to accept any Sakigake awards.

■ Also included in the budget of the Science and Technology Agency (STA) is a surprising request of ¥488 million (\$3.6 million) for research on development of a large solid-fuel rocket in collaboration with the academic Institute of Space and Astronautical Science.

It was only a few weeks ago that the agency submitted a proposal to the Space Activities Commission for joint development of the rocket (*Nature* 352, 181; 18 July 1991). The decision by the commission and the institute to accept the proposal has been much faster than most observers expected, allowing the agency to submit a budget request this year.

The J-1 rocket, as it is to be called, will combine the space institute's M-3SII rocket with the solid-fuel booster rocket of the H-II rocket, being developed by the STA's National Space Development Agency. J-1 will be capable of lifting medium-sized, one-ton satellites into low Earth orbit at very low cost.

David Swinbanks

British Science

We shall be publishing on 12 September *Nature's* Manifesto for British Science, which will be offered without charge to any one of the political parties contesting the next general election, and in the belief that it is an improvement on their own version.

Following publication, on Friday 13 September, there will be a public meeting at the Royal Institution, Albemarle Street, London W1 at which Sir Mark Richmond, chairman of the Science & Engineering Research Council, will be the principal speaker. The chairman will be John Maddox, Editor of *Nature*. The meeting will start at 9.45 a.m. and end at 12.45 p.m. The political parties will be invited to send representatives, but there will be ample time for discussion. Those wishing to attend should apply for tickets to Mary Sheehan, *Nature*, 4 Little Essex Street, London WC2R 3LF.

BUDGET REQUEST FOR SCIENCE AND TECHNOLOGY AGENCY

	thousand million yen	% change from 1991
Total R & D budget	556.1	+ 6.4
Special Promotion Funds	11.1	+ 5.7
Space	144.9	+ 10.0
Nuclear Energy	317.3	+ 3.5
SOR (Spring-8)	7.2	+ 47.1
Ocean Research	11.5	+ 7.6
ERATO	6.4	+ 12.3
Human Frontier Science Program	2.5*	+ 13.6
Green Planet Project	19.6	+ 90.3
Sakigake Research	1.2	+140.0
Inspiration	0.3	(new)
Human Genome	1.5	+66.7

* Ministry of International Trade and Industry has requested ¥1.7 thousand million for Frontiers, bringing total request to ¥4.2 thousand million.

Small wonders in Africa

Paris

DESPITE a year of government investigations into his controversial AIDS vaccine studies in Africa, French researcher Daniel Zagury is planning yet another expedition to that continent. But his principal equipment this time will more probably be a ruler than a syringe. With Arsène Burny of the University of Brussels, Zagury intends to explore a mysterious correlation between short stature and resistance to AIDS. The correlation was discovered two years ago by Belgian scientist Philippe Van de Per and his colleagues in Rwanda in a group of about 700 children infected at birth by their HIV-positive mothers.

Most seropositive babies die in their second or third year. But 14 of these Rwandan children have shown a remarkable ability to survive; all are more than four years old, and some have reached the age of 14 with relatively good health. If Zagury's hypothesis is correct, these children can thank an excess of the neuropep-

that." Testing the blood of the children confirmed unusually high levels of somostatin and, as a result, low levels of growth hormone.

Robert Gallo, a US AIDS researcher at the National Institutes of Health, has seen Zagury's data and says the implications are "very interesting". But he is reluctant to draw conclusions until Zagury and Burny get more samples from a larger group of seropositive children, something the researchers plan to do next year. "There is definitely a correlation with stature [and

delayed onset of AIDS] in their data," Gallo says, "but I don't know how far to push it yet." Nevertheless, he describes the discovery as "a startling lead".

The children are not entirely healthy; their CD4 cell count is only in the range of 200 to 400 — about one-third of normal levels. But that is still well over the cell count that a person with full-blown AIDS might have. And the children are still alive, something that Zagury finds most persuasive of all. If somostatin can indeed suppress the effects of HIV, he argues, science could have a powerful tool in the struggle to keep AIDS patients alive.

Christopher Anderson

CLINICAL TRAILS

Blazing an ethical trail

Washington

A CLINICAL trial to examine whether the AIDS drug zidovudine (AZT) can reduce the transmission of human immunodeficiency virus (HIV) from pregnant women to their fetuses is now getting under way, having tiptoed its way through an ethical mine field on the way to approval. As one of the few major trials of a drug known to have toxic side-effects to focus on pregnant women, the new study has broken new ethical ground, and may have cleared an easier path for subsequent studies looking for ways to prevent the transmission of HIV from infected mother to fetus.

The trial is still opposed by a number of AIDS advocacy groups, who argue that the study neglects the health needs of the HIV-infected women themselves, but such criticism from activists is not uncommon. What is uncommon is the intensity of the ethical debate that the trial triggered among medical researchers, pitting paediatricians against AIDS researchers.

The proposed study was predicated on the assumption that if the number of viral particles circulating in the mothers' blood can be reduced, then transmission of HIV to their fetuses (reckoned to occur in up to 40 per cent of pregnancies when the mother carries the virus) will be less likely. This may be a reasonable assumption, but there is no animal model in which to test it.

Deborah Cotton, from the Harvard School of Public Health and a member of the Food and Drug Administration (FDA) Antiviral Advisory Committee that reviewed the study, says she was initially "aghast" when she learned of the plan to launch a trial giving AZT to pregnant women. AZT can have toxic side-effects, including anaemia, severe nausea and muscle wasting, which force many AIDS patients to discontinue the drug.

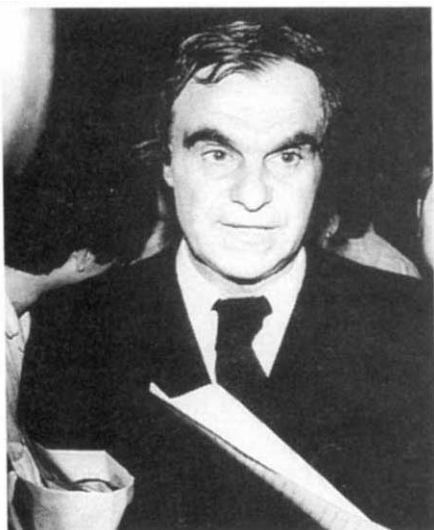
Catherine Wilfert, a Duke University paediatrician specializing in AIDS who was called in by the FDA to provide expert advice on the trial, says the main difficulty was reconciling the different ethical per-

spectives of paediatricians and mainstream AIDS researchers. The controversy centred on the dangers of exposing fetuses to a highly toxic drug, particularly for the majority of babies born to HIV-infected women that would not, in any case, become infected. Somewhat surprisingly, the loudest protests came from AIDS researchers working with adults. This was in large part due to the tendency for these researchers to view AZT as a treatment for HIV infection, and one that should not be applied when not strictly necessary, because of its toxic effects. Wilfert says the change of heart came once the mainstream AIDS researchers became convinced that the trial should be seen as one looking at the prevention of HIV infection, rather than its treatment.

After more than two years of soul-searching and stringent review (the FDA antiviral committee is usually not asked to examine drug trials while they are still being planned), the trial was eventually approved — and with the protocol rewritten to give AZT during both the second and third trimesters of pregnancy, rather than just the third as in the original draft.

Having cleared its ethical hurdles, however, the trial now faces new obstacles. Recruitment has so far been slow — the National Institute of Allergy and Infectious Diseases (NIAID), which is coordinating the trial, hopes to include some 750 women, but since recruitment began in April, only 36 have enrolled. The main problem seems to be that HIV-infected pregnant women, many of whom are drug users, typically come from a section of society in which participation in basic health care programmes, let alone in clinical trials, is poor. "These women are about as disadvantaged as they come," Cotton observes. Given this constraint, James Balsley, medical officer for the study at NIAID, says it could take three years to enrol the full complement of 700-plus women.

Peter Aldhous



Zagury finds advantages in shortness

tide somostatin for their longevity. Just as somostatin inhibits the growth hormone somatotropin in the endocrine system, Zagury believes it may suppress the damaging effect of the human immunodeficiency virus (HIV) on the immune system. That, he argues, may have delayed the onset of AIDS in the children.

Contrary to rumours circulating in Parisian laboratories about the discovery, Zagury says, the AIDS-free children "are not pygmies, they're just very short." He says that although he had suspected a link between the growth regulatory system and the immune system, it was not until he and Burny actually explored the health records of the small group of healthy children that they found some confirmation. "We immediately saw that they share a common clinical denominator — their size. They are not dwarfs, but they're just above

California tackles insurance

Washington

THE California state legislature is poised to pass a bill that would place an eight-year ban on the use of genetic information to discriminate among people in selling health insurance. If the bill is not vetoed by Governor Pete Wilson, California will become the first state to prevent health insurers from using the results of genetic tests to influence underwriting decisions.

The bill, sponsored by Lloyd Connelly, a California Democrat, would also ban the use of genetic test information in employment decisions. The original draft proposed an indefinite ban on health insurers using genetic test results, but in a compromise to reduce opposition from the insurance industry, the bill emerged from the state Senate Judiciary Committee last month with the health insurance ban curtailed to an eight-year moratorium.

The bill is now expected to be passed by the legislature within a matter of days, and then needs only the governor's signature to become law. Wilson's aides say he has not yet taken a position on the bill.

Coming from the most populous state in the country, legislation in California is watched carefully — though not necessarily imitated — by other states. Given that the Human Genome Project is likely to yield many new diagnostic tests for genetic diseases over the next decade, Steve Brown from the Council of State Governments says he expects many states to regulate the use of genetic information over the next few years. But he says it is difficult to predict how other states will act on the controversial issue of health insurance.

Insurance companies argue that genetic test results are no different from the other medical data they use in underwriting health insurance. The industry fears that if applicants for health insurance are able to withhold genetic information, those whose genetic test results reveal that they are likely to have high health care costs will seek out health insurance in greater numbers. The result would eventually be higher outlays for insurance companies, and the undermining of the industry, insurers say. Nevertheless, Brad Wenger, from the Association of California Life Insurance Companies (which also represents the state's health insurers), says his organization has dropped its opposition to Connelly's bill because most companies do not expect to use genetic information on a large scale over the next eight years. By the end of that period, he says, the federal government will probably have passed legislation on the issue.

On the other hand, some medical geneticists argue that genetic information is fundamentally different from the other data used by health insurers. If tests become available to identify a wide range of

genes that confer susceptibility to particular health problems, they say, the whole notion of 'shared risk' that underlies the insurance business will be subverted, and people whose genetic makeup indicates a higher probability of illness will find it difficult to get affordable health insurance. (See also page 2.)

Paul Billings, a geneticist at the California Pacific Medical Center who helped Connelly to draft his bill, believes legislation is needed now because, he says, some companies are already beginning to discriminate unfairly against people with genetic disorders. For example, he says,

SCIENTIFIC SOCIETIES

Ethics rules provoke storm

Washington

D. ALLAN Bromley, science adviser to President George Bush, may soon be asked by many of the leading US scientific societies to secure changes to new government ethics rules that the societies say could undermine their very existence. At a meeting in Washington last week, representatives from more than 15 scientific societies decided to seek a meeting with Bromley to explain their opposition to proposed rules that could, they say, prevent government researchers from serving as society officers.

In late July, the Office of Government Ethics proposed a wide-ranging revision of its ethical standards for all employees of the US government's executive branch. Most of the rules are uncontroversial — for example, a prohibition on government employees receiving gifts from people with whom they do business — but the new section on 'outside activities' has sent a frisson of panic through the science societies. According to some interpretations, the rules would prohibit government scientists from doing almost any society work on government time. The societies fear the rules would demand that government researchers editing journals or organizing meetings on behalf of their societies should do this in their spare time.

The societies argue that the proposed rules would make government researchers into the 'second class citizens' of the scientific community, unable to participate fully in their professional organizations. Worse still, universities and industrial employers might follow the government example and also place restrictions on researchers working for scientific societies during work hours.

One attorney at the Office of Government Ethics explains that the rules are designed to prevent society officers from using government time to conduct routine society business, such as contacting mem-

bers to remind them to pay their subscriptions. But the rule would allow a government researcher to attend a society conference if the subject is directly related to the mission of the researcher's agency. As to whether government scientists would be allowed to organize a conference or edit a society journal, the attorney says these are "grey areas" that will need further thought.

Many science society officials argue that the proposed rules would also hurt the government. The progress of science depends partly on the existence of societies to promote conferences and publish journals, they argue, and if the societies are damaged by the new rules, then the government, as the largest funder of science in the United States, will also suffer.

Stressing the disadvantages to the government if its researchers are prevented from working for their professional societies may be the best way to attack the proposed rules, but the coalition of scientific societies faces a difficult dilemma: how to exert maximum influence within the US Administration, while attracting a minimum of press attention. In the wake of a series of well-publicized scientific misconduct investigations, many journalists are now primed to view scientists as cheats. So any suggestion that researchers are seeking to use government time to further the ends of their professional societies is unlikely to attract favourable coverage.

The delegation to Bromley, assuming that he agrees to the meeting, is expected to consist of four or five prominent society presidents. But for many of the societies gathered at last week's meeting, the immediate priority is to obtain an extension to the period allowed for public comment on the rules. Unless the Office of Government Ethics allows an extension, the societies have only until 20 September to produce their responses.

Peter Aldhous

US national labs face changing roles

Washington

THE three US nuclear weapons laboratories — Los Alamos, Lawrence Livermore and Sandia — can rest easy now, at least for a while. Last month, President George Bush signed into law a 1992 budget for the labs which avoided proposed cuts that would have, according to the lab directors, jeopardized their ability to perform their basic job: to develop state-of-the-art nuclear warheads.

But the debate over the laboratories' future is far from over. As the perceived nuclear threat from the Soviet Union continues to diminish and as pressures mount to trim the federal deficit, budget makers are likely to view the laboratories as one place where cuts can be made. In response, the laboratories and the Department of Energy (DOE), which runs them, are taking a new look at what the laboratories' national security role should be in this changing world.

In a few weeks, the DOE will release a report from a special task force on the national laboratories that addresses this question. The report will recommend that the laboratories apply the expertise of their 9,000 scientists and engineers who develop nuclear weapons in new and different ways, from designing a new generation of safer warheads to assisting in verification of arms control treaties and helping in the development of a smaller, environmentally friendly nuclear weapons production complex. The waning of the Cold War is modifying what the nation asks of its nuclear weapons scientists, and the change is coming fast.

Clean-up squeezes budget

Historically, the budget for nuclear weapons work at the laboratories has moved up and down with the tide of nuclear politics. The most recent high tide came in 1987, after the Strategic Defense Initiative had pumped hundreds of millions of dollars into the laboratories' core nuclear weapons work, referred to as Research, Development & Technology (RD&T). Since then, according to Tony Lane, director of DOE's office of programme analysis and financial management, a 12 per cent decline in funding has forced the laboratories to cut their RD&T workforce from 11,150 positions in 1987 to 8,900 estimated for 1992.

And it could have been worse. If the budget submitted to the Office of Management and Budget in the autumn of 1990 had passed, another 500 RD&T workers would have lost their jobs, says Richard Claytor, assistant secretary of defence programmes at DOE. But Sen. Pete Domenici — who represents New Mexico, home to Los Alamos and Sandia — persuaded the Senate Appropriations

Committee to give the DOE another \$200 million for its nuclear weapons work, bringing the total RD&T budget to \$1,945 million.

Without that \$200 million injection, the laboratories' baseline competence in developing nuclear weapons would have been threatened, Claytor says. "We are now approaching a point where further reductions in the population of nuclear specialists can be harmful."

Nonetheless, DOE itself had submitted the original budget that department and national lab officials now say would have menaced nuclear weapons competence, and the reasons for that decision offer a telling insight into the conflicting pressures the laboratories are facing.

While RD&T budgets have been dropping, the DOE's environmental budget — propelled by the necessity of cleaning up the nuclear weapons manufacturing complex — has been swelling. In just two years, from 1990 to 1992, DOE's budget for waste management and environmental restoration doubled from \$2,300 million to \$4,600 million. At the same time, congressional budget agreements aimed at bringing the federal deficit under control have put a cap on the total annual DOE budgets. Because many of the environmental cleanups are mandated by law, DOE is forced to find cuts elsewhere in its budget, and RD&T was one of them for fiscal year 1992.

Domenici, supported by DOE and laboratory officials, argued that the RD&T cuts would be debilitating to the laboratories' nuclear design capabilities, and he convinced the Pentagon and Congress to agree to transfer \$200 million from the Department of Defense budget to DOE's nuclear weapons budget.

Nevertheless, the national laboratories' nuclear weapons work is likely to continue to be squeezed between the budget cap and the growing environmental expenses. And it is a catch-22 in more ways than one. One congressional staff member familiar with the situation characterized it as a power struggle between congressional committees: "The environmental committees are fighting the armed services committees [which oversee the nuclear weapons budgets] for jurisdiction over cleanup. No one on the armed services committees wants the environmental committees to have authority over the weapons plants." But the price of protecting the nuclear weapons complex from the environmental committees is to watch cleanup costs eat up an increasingly large part of the DOE budget.

Meanwhile, officials at DOE and the nuclear weapons laboratories are attempting to redefine what the laboratories' roles should be in a world that is less worried

about nuclear war. All three of the laboratories have assignments unrelated to nuclear weapons development, but the weapons work is the core and the *raison d'être* for each of them and accounts for approximately half of each of their budgets.

New directions

When the Secretary of Energy Advisory Board, a panel formed of prominent non-DOE scientists to counsel Secretary of Energy James Watkins, releases an interim report on the role of the laboratories in the next few weeks, it is likely to recommend expanded roles for them in several areas. The recommendations, which were discussed briefly in a public hearing by the board in July, will include:

- Researchers at Los Alamos, Livermore and Sandia should play a major role in the clean-up of the nuclear weapons production complex. This will continue to be a growing part of DOE's job, and the nuclear weapons laboratories — with their expertise in nuclear physics and radioactivity — are natural places to do much of the research.

- The laboratories should help in developing a new nuclear weapons production complex. The DOE has set a goal of achieving a new complex that is more compact, less expensive and generates less nuclear waste than the current decades-old complex, but Siegfried Hecker, director of Los Alamos National Laboratory, notes that with the departure of such long-time contractors as Du Pont and Dow, the DOE has lost much of the expertise in the production complex. This leaves the laboratories as the natural — and sometimes the only — recourse. "Where else are you going to go for plutonium processing R&D?" Hecker asks.

- The issues of nuclear non-proliferation, nuclear arms control and treaty verification are likely to take on increasing importance in coming years, and the laboratories — which already do some research on the technical aspects of these matters — are again the natural place to spend the extra money.

- The laboratories should turn their nuclear weapon design efforts toward making nuclear weapons safer — less likely to release radioactive materials in case of such accidents as fires or plane crashes. Weapons designers say the warheads are already quite safe, but as the likelihood decreases that they will ever be used, the safety of the weapons becomes more important relative to their explosive power, and these changing priorities argue for redesigned warheads.

One issue that is likely to come up again and again is: Is the present duplication in the laboratories' nuclear design

BIOTECHNOLOGY

efforts justified? As those efforts are now set up, both Los Alamos and Lawrence Livermore perform basic design work, while Sandia has a support role — putting together the electronics and other parts of bombs so that they work as planned. Some argue that the country does not need two nuclear weapons laboratories that are doing the same job.

Checks and balances

DOE and laboratory officials disagree vigorously. "Having two design labs that can balance each other and check each other is very important," says Claytor at DOE. The two laboratories have a history of approaching problems very differently (which is often traced back to the differences and competition between the founding figures Robert Oppenheimer at Los Alamos and Edward Teller at Livermore), and these varying approaches and competitive spirit provide an effective method of peer review in a field that otherwise would have none. "As the number of nuclear tests decreases, the country still has to answer questions about its nuclear forces," says John Nuckols, director of Lawrence Livermore, but with only one lab, there would be no way to settle disputes over weapons design.

The question of whether to pay for two laboratories really depends on how good the United States wants its nuclear arsenal to be, notes Roy Woodruff, formerly associate director for Nuclear Design at Livermore and now a staff scientist at Los Alamos. "If your national policy is minimal deterrence, then one lab may be enough. If you're trying to keep a competitive edge over a very capable adversary, you need something more."

Although Congress answered that question last month by voting the extra \$200 million to keep nuclear design work healthy at both laboratories, this is unlikely to be the final word. Claytor says he foresees "stability" in the RD&T budget in coming years, but this is going to be difficult to obtain unless the growing environmental costs level off.

Woodruff says that he, like many nuclear weapons designers, has always hoped that eventually he would work himself out of a job — that nuclear deterrence would hold off war long enough for peace to be kept by other methods, and nuclear weapons would not be needed. Although this seems unlikely to happen soon, still an increasing number of weapons designers could find themselves out of work as the nation decides it has less need to continuously upgrade its arsenal. And one way or another, the laboratories are in for a decade of unprecedented change, and their new courses are going to be set by decisions made in the next couple of years.

Robert Pool

Will milk shake up industry?

Washington

THE times they are a-changing, down on the farm. Ever since researchers demonstrated the tissue-specific expression of several types of foreign proteins in the milk of transgenic mice, the race has been on to scale up the technology to dairy livestock, where the capacity for high-volume production has obvious appeal.

Now in a significant technological advance, three separate research groups from the United Kingdom, the United States and the Netherlands have made considerable progress in this regard in sheep, goats and dairy cows. The research, reported in the September issue of *Bio/Technology*,



brings the biotechnology industry one step closer to the goal of providing a cheaper means of producing proteins than conventional expression systems.

Success with larger animals was first reported in 1989 by Agriculture and Food Research Council (AFRC) researchers at the Institute of Animal Physiology and Genetics Research in Edinburgh, Scotland. Although the Edinburgh team produced transgenic sheep that secreted human antithrombin factor IX and human α_1 antitrypsin in their milk, expression levels were low.

Building on this, a joint project between Alan Coleman and colleagues at Pharmaceutical Proteins Limited and researchers at the Edinburgh institute produced five transgenic sheep — four females and one male — into which the human α_1 antitrypsin gene had been incorporated. Following mating and breeding of three of the females, milk from all three contained human α_1 antitrypsin at levels of greater than 1 gram per litre, and, in one case, levels of up to 35 grams of human α_1 antitrypsin.

Researchers at Tufts University in Massachusetts and at the Massachusetts-based biotechnology company Genzyme Corporation have developed the first

transgenic goats to produce tissue plasminogen activator (tPA) in their milk. Although the level of production — about 3 micrograms of tPA per millilitre of milk — is not economically realistic, Karl Ebert, director of experimental biotechnology at Tufts, says he has recently produced another transgenic goat in which the levels were increased a thousand-fold.

From a commercial standpoint, the dairy cow is the animal of choice. Dairy cows can produce 10,000 litres of milk annually, as compared to sheep and goats, whose milk yields can vary from 250 to 800 litres a year, depending on the breed. But historically, efforts to produce transgenic dairy cows have been thwarted because of cumbersome and costly surgical procedures. Now, however, researchers from Gene Pharming Europe (Leiden), the University of Leiden and the Research Institute for Animal Production (Zeist) have circumvented the need for the surgical removal and transfer of embryos by combining gene transfer with an *in vitro* embryo production system.

Jonathan MacQuitty, president and chief executive officer of GenPharm International, Inc., Gene Pharming's parent company, says GenPharm plans to use the technology to produce hLF, a human milk protein with antibacterial and iron transport properties, for infant formula preparations and as an oral treatment for immunocompromised patients. The company also intends to produce other proteins, which, like hLF, would be difficult or impossible to produce efficiently by other means.

Diane Gershon

TELECOMMUNICATIONS

Pandolfi boosts HDTV

London

IN hopes of encouraging European electronics companies and broadcasters to support new High Definition Television (HDTV) technology, the European Communities (EC) are planning to double their subsidies next year, EC minister for research and telecommunications Filippo Pandolfi announced last week. In a new five-year HDTV strategy that is to begin in 1992, the EC will spend a total of 1,000 million ecu (\$1,200 million) — 200 million ecu per year — on the broadcasting industry in hopes that a successful launch of HDTV by 1994 will rescue the fading European television and electronics industry. But Pandolfi has an uphill battle for European acceptance of HDTV; a new study by the accounting firm Coopers and Lybrand finds that switching to the new standard will cost consumers 21,000 million ecu (\$25,000 million) in the next 10 years to upgrade their receiving equipment.

Christopher Anderson

EC look to animal patents

London

Six years after a patent application for the genetically engineered 'Harvard mouse' was submitted to the European Patent Office — and two years after the same patent was granted in the United States — the application still languishes in a legal quagmire of moral and procedural debate over the patenting of life forms in Europe. But a decision next month in Brussels may mean that the end is in sight. In the session beginning on 7 October, the European Parliament legal affairs committee is expected to approve a draft directive "on the legal protection of biotechnological inventions" that would explicitly allow the patenting of bioengineered life forms — everything from a gene to an entire animal — in the nations of the European Communities (EC).

If the directive is indeed approved, it will be voted on by the full Parliament, and then go to the Council of Ministers, which will make a final decision based on the parliamentary recommendation and the proportions of the vote. But before the directive emerges, legislators must navigate their way through a hail of ethical and legal arguments from outspoken opponents of animal patenting.

Unlike countries such as the United States, Canada and Hungary, where patents for plants are permitted, Europe's current Patent Convention does not generally allow for property rights over plants and animals. Any exception requires a protracted legal battle — the Harvard mouse imbroglio, for example — or new legislation, as is now being considered in Brussels.

Both the proposed EC legislation and the current intellectual property rights negotiations at the General Agreement on Tariffs and Trade (GATT) world trade talks in Geneva seek to bring US-style

biotechnology protection to the rest of the world. In favour of these efforts, industry organizations such as the London-based BioIndustry Association argue that Europe's biotechnology industry will continue to lag behind that of the United States until its patent laws are equally strong. "Both investment and the progress of research will be concentrated in those countries that offer the best incentives, especially patent protection," says executive director Louis De Gama.

But opponents of the legislation are loud and organized. They argue that patents for genetically engineered plants and animals will decrease genetic diversity, slow the free exchange of scientific knowledge, raise ethical problems such as the morality of monopoly ownership of living things, and increase animal suffering.

One of the strongest voices of dissent comes from the developing nations, which argue that much of the raw genetic stock of agricultural biotechnology — genes from disease- and pest-resistant plants that can be inserted into other species — have come from their natural resources. As Western companies get rich from new bioengineered plants, they say, the work of the West African farmers who raised the original insect-resistant cowpeas and the Bolivian farmers who bred disease-free potatoes is being looted. "Genetic resources must be considered as our common shared heritage" said Ed Mayo of the World Development Movement last spring, when more than 30 European environmental, animal welfare and agricultural groups organized by the London-based Genetics Forum urged the rejection of the directive. Both the Genetics Forum and the BioIndustry Association are preparing to launch another lobbying campaign this month to shift the EC debate their own way. **Christopher Anderson**

TELECOMMUNICATIONS

HDTV satellite is up

Tokyo

OPERATORS of Japan's satellite television breathed a sigh of relief Sunday after the successful launch of a new broadcasting satellite by the National Space Development Agency. The satellite has a channel to broadcast High Definition Television (HDTV) and will also provide vital backup for a companion satellite that is limping along on reduced power.

Satellite television has been on the brink of collapse in Japan after two broadcasting satellites — one aboard a European Ariane rocket, the other on a US craft — were blown up shortly after lift-off due to malfunctions of the rockets. Adding to these woes was a malfunction of the BS-

3a satellite launched by the space agency last year, which has only about 75 per cent power because of a defect in its solar panels.

In addition to backing up the BS-3a, the new satellite will enable Japan's national broadcasting company to extend test transmissions of HDTV from about one to eight hours a day, and Japan Satellite Broadcasting Inc. will introduce the world's first commercial HDTV transmissions with the satellite — all of which should help nurture Japan's fledgling HDTV industry. But with HDTV sets currently priced at about \$30,000 apiece, no one expects a rush of new customers yet.

David Swinbanks

First the moon, now Mars

Tokyo

THERE seems to be no limit to the ambitions of Japan's little academic space institute, the Institute of Space and Astronautical Science (ISAS). Having sent a probe to the Moon last year, ISAS scientists are now planning to stretch their technology to its limits and send a modest mission to Mars in 1996.

Earlier ISAS plans had been more conservative and had called for a mission to Venus after a follow-up expedition to the Moon in 1995 (*Nature* 344, 693; 1990). At the time, Mars was considered to be beyond the range of the institute's next-generation M-V rocket, which is currently under development.

But ISAS scientists now believe they can reach Mars with the M-V by using a smaller probe with its spin axis and communication antenna and solar panels constantly pointed towards the Earth. A fortuitous alignment of the Sun, Earth and Mars will allow the probe to catch sufficient solar power without having to align the solar panels directly towards the Sun and without having to install heavy steering equipment for the communication antenna.

The saving in weight will amount to 50 kg, bringing the total probe weight down to 250 kg (excluding fuel) — light enough that the little M-V solid fuel rocket will be able to push the probe all the way to Mars. But the light-weight probe will only be able to carry a very small scientific payload of about 40 kg. In contrast, the Soviet Union is planning to send a 2-ton payload to Mars in 1994 — nearly 50 times as heavy as the Japanese probe will carry — and the United States will send a probe with about a 1-ton payload next year.

Koichiro Tsuruda of ISAS says Venus and Mars are equally interesting for scientists, but ISAS has chosen Mars as the first planetary target because its upper atmosphere is less studied than that of Venus. If all goes according to plan, the Mars probe, called Planet B, should arrive at Mars in September 1997 and will orbit the planet at distances ranging from 150 to 34,000 km. The main scientific objective will be to study the Martian upper atmosphere and its interaction with the Solar wind, Tsuruda says.

The light-weight Mars probe can be built for only about \$80 million, with an additional \$40 million or so for the M-V launcher, Tsuruda says. In contrast, the US and Soviet Mars missions are expected to cost hundreds of millions of dollars. The Ministry of Education, Science and Culture has just requested about ¥100 million (\$740,000) for Planet B in the budget for next fiscal year.

David Swinbanks

Open letter to Paul Doty

DEAR PAUL — Your recent comments on my activities (*Nature* 352, 183; 1991) require a brief response.

The standards of scientific investigation are the highest standards to which humans can aspire. Trying to understand the secrets hidden away in nature is a difficult and trying enterprise that requires a critical assessment of each piece of evidence and a sceptical review of every idea. The judgement of past human activities is even more demanding because there is no possibility of repetition.

You have used a leaked, confidential draft of a government report as a basis for definitive judgements about the acts of others. Neither you nor those judged have had full access to the evidence. Your judgements therefore could not have derived from the careful assessments that you yourself say that scientific evidence requires. The staff of the Office of Scientific Integrity has recognized that a fair judgement requires access to the evidence. But your judgements do not depend on complete evidence; your verdicts are, rather, based mainly on the unsubstantiated, and often refuted, allegations of one participant in events five years old.

You say that I failed traditional standards of science but you have not discussed the events with me, choosing to rely on information from others. On that basis you have enlisted in a campaign of criticism. Paul, there are a few things you should know and I would like to share.

I believe that my science — including the Weaver *et al.* paper — is done with rigour and criticality. But, because the issue is one of judging historical events, the only way to judge that statement is by the traditional test of science: have the data proved reliable?

For the Weaver *et al.* paper, the data have proved more durable than the data in most papers. No experiments of which I am aware have appeared in the literature that contradict the data of the Weaver *et al.* paper. In fact, there is much published evidence* and more coming that support the paper's results in remarkable detail.

One of your statements, Paul, is unworthy of you. You charge that an "embarrassing" band was eliminated from a figure by underexposure of the image. That is plain wrong: no one consciously eliminated any band. There was a faint band that we saw in the original autoradiogram that was not visible in the published figure. That loss occurred at some indeterminate time in the publication process, as has happened to many faint bands. In any case, it was

not embarrassing; its existence was not of significance to the paper, otherwise it would have been investigated further. Even if that band could be shown to represent a small expression of μ RNA in that particular hybridoma, the interpretation that the cell was making a gamma protein with idiotype characteristics and a V region like that of the transgene would not be changed. That was the issue at stake there. I urge you to go back and study this again.

I welcome your call for the highest standards in the prosecution of scientific investigations. I have supported just those values all of my scientific life and will continue to do so.

In contrast to your apparent certainty about the events of the past, I am not now convinced that I know all the answers. But I do know that I have tried to ferret out the truth and I feel that I did reasonably well because the science has stood up to the toughest test of all, the test of history.

DAVID BALTIMORE

Rockefeller University,
1230 York Avenue,
New York,
New York 10021-6399, USA

*J. clin. Invest. 79, 1044–1053 (1987); Eur. J. Immunol. 17, 399–403 (1987); Proc. natn. Acad. Sci. U.S.A. 85, 3546–3550 (1988); Proc. natn. Acad. Sci. U.S.A. 86, 2346–2350 (1989); Int. Immunol. 3, 67–73 (1991); Int. Immunol. 3, 185–196 (1991).

The case against

SIR — Your leading article, "The case for the human genome" (*Nature* 352, 11; 1991), shows breathtaking ignorance of some historical and ethical aspects, which will be used as arguments against the genome project. The statement "Those who in the 1930s mounted the holocaust hardly bothered to excuse themselves by reference to genetics, although the ill defined and empty doctrine of Aryan superiority may have stilled the consciences of some practitioners," is a travesty of the truth and cannot pass unchallenged. Are you not aware that it was the geneticists and related scientists of the time, including Fischer, Von Verschuer and Lenz, who provided and supported the supposed scientific foundations on which the whole philosophy was based¹, and who supported its putting into practice along with their medical colleagues? It is hardly surprising that the article expresses bewilderment with present-day attitudes in Germany that are attempting to reconcile future progress with past abuses.

In dismissing the likely problems of genetic testing in relation to insurance, the leading article states: "the contention that people unlucky enough to carry

identifiable genetic abnormalities should not be denied insurance on the same terms as other people begs the question why people relatively free from identifiable genetic abnormality should therefore pay more than would otherwise be necessary". Such a view totally ignores the complexity of issues involved, including third party pressures for testing for late-onset disorders, where serious problems are already arising.

Most responsible scientists are keenly aware of the ethical problems that their work may raise and are glad to see them debated and, where possible, resolved. To suggest that the ethical problems raised by the human genome project are "insubstantial" and should be "handled delicately", is to give support to those in the community who contend that scientists cannot be trusted to act responsibly on ethical issues.

PETER S. HARPER

University Hospital of Wales,
Heath Park, Cardiff CF4 4XW, UK

1. Weindling, P. *Health, Race and German Politics between National Unification and Nazism, 1870–1945* (Cambridge University Press, 1989).

On the run

SIR — Where have the Arizonans Ronald Dorn and James Clark (*Nature* 352, 10; 1991) been for the past century? Not, I trust, on the run. To find an answer to their question about the use of the term 'run' to connote a session on an instrument, one need look no further than the *Oxford Dictionary*. Among numerous definitions of 'run' one finds "spell of making or allowing machinery to run or continue to work", with the year 1875 given as the first documented date of the use of this particular meaning. Scientific analysis is probably just a relative newcomer to this long established terminology.

DAVID WEITZMAN

Cardiff Institute of Higher Education,
Cardiff CF5 2SG UK

Help wanted

SIR — The Committee on Medical Aspects of Food Policy of the Department of Health has set up a Working Group to review the nutrition of young children during the weaning period. The working group is concerned that its deliberations should encompass all opinions in its field and invites concise written submissions based on reasoned argument and scientific data from interested parties. Submissions should be sent to me.

P. CLARKE
(Nutrition Unit)

Department of Health,
Room 541, Wellington House,
133–155 Waterloo Road,
London SE1 8UG, UK

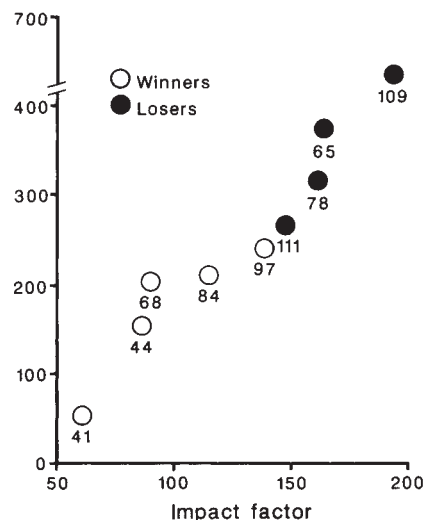
Academic promotion in Italy

SIR — In Italy, appointments to tenured positions at the full professor level are not handled through local *ad hoc* search committees or advertisements in international scientific journals; instead, all vacancies are submitted to the Ministry of University and Scientific Research and Technology (MURST) in Rome, which every four to five years announces a national competition for all disciplines.

This peculiar procedure contributes adversely to the precarious state of health of academic institutions in Italy. For example, in 1988, among the several hundred tenured positions vacant in schools of medicine, five were for professorships in a specialty branch of internal medicine. A committee of five professors in that specialty, with one acting as chairman, evaluated 16 applicants. The committees are asked by state law to ascertain only the "full scientific maturity" of applicants, through examination of their scientific output; there are no other requirements or specific guidelines and the candidates are never interviewed. The discretionary powers of a committee are almost unlimited. It is therefore possible for candidates for a tenured position involving heavy responsibilities and duties in teaching, research and clinical practice to be evaluated without taking into consideration the applicants' previous experience in teaching, clinical skills and proficiency in grant applications. A final decision about the best five was finally reached in November 1989. The committee's proceedings were available after several months, but there was no mention of the criteria used to evaluate the applicants' "scientific maturity".

We therefore decided to check whether widely applied indicators for evaluation of research based on bibliometric analysis did support the choices of that committee. A search was made on the five successful candidates (winners), and on four rejected applicants (losers) who are especially well thought of in the scientific community.

The MEDLINE database provided the list of all publications from 1966 to 1988. However, a publication count gives only a measure of the total volume of the research output and does not indicate the quality of the work. Because citation frequency analysis has been increasingly used in the evaluation of science, a search was made of the citations for each candidate. Citations from 1965 to 1988 were obtained from the *Science Citation Index* (SCI) database for the articles where the candidates appeared as first or last author. This approach overcomes, at least partly, the criticism that in a routine search citations are awarded to



Correlation between cumulative impact factor and number of citations. Under each circle the total number of papers is reported.

the first author. The mean of the citations for the four excluded candidates was 397 ± 165 , versus 172 ± 74 for the winners ($P < 0.01$).

Another reliable estimate of the performance of scientific papers is their SCI impact factor (IF). The IF is the ratio of the number of citations a journal receives to the number of papers published over a period of time. Furthermore, a paper published in a high-impact journal indicates careful peer review even for controversial ideas. All publications

were then rated according to the 1988 IF. The four losers had more articles published in the 15 most-cited journals (15 vs 4). The five winners had a cumulative IF score largely inferior to that of the four losers; the difference between the means of the two groups (99 ± 30 versus 169 ± 19) was statistically significant ($P < 0.005$). A positive correlation between cumulative IF and citations was observed (see figure). On the basis of our studies, it is therefore unlikely that the four rejected candidates were deemed unsuitable for an academic promotion to full professor because of insufficient "scientific maturity". These international scientific indicators were ignored and the scale of values adopted by the committee is a mystery. However, the names of one or more members of the committee itself often appeared among the multiauthored papers of the winners (for one of the successful candidates this happened in 70 per cent of the articles).

We conclude that the present machinery for academic promotion, besides being artefactual and unreliable, is very much like an 'old boy' network, and has led many of Italy's leading researchers to leave the country. There is an urgent need for new rules and guidelines to be incorporated into the decision-making process of academic promotion.

GIAN FRANCO GAETANI

ANNA MARIA FERRARIS

University of Genova Medical School,
ISMI-Viale Benedetto XV, 6
16132 Genova, Italy

Biological accord

SIR — Lennart Philipson's Commentary, "Turmoil in European biology" (*Nature* 351, 91; 1991), was a timely analysis of the dire straits through which basic biological research in Europe is passing — with little sign of calm water and prosperity ahead.

Part of Philipson's remedy is the formation of a joint coordinating body for all organizations representing biological research in Europe. As it happens, plans to establish such an organization had been under discussion for more than two years, following an initiative taken by Professor Karl Ullrich and Mr B. Sälzer, MEP, in 1989, and EUSEB (the European Union of Societies for Experimental Biology) officially came into being in Würzburg, Germany, on 1 June 1991. The founder members of EUSEB are the European Cell Biology Organization (ECBO), European Developmental Biology Organization (EDBO), European Neuroscience Association (ENA), European Society of Toxicology (EUROTOX), Federation of European Biochemical Societies (FEBS), Federa-

tion of European Pharmacological Societies (EUROPHARM) and Federation of European Physiological Societies (FEPS), with the European Molecular Biology Laboratory (EMBL) as an associate member.

Because, as far as we are aware, no registry of all the federations of European biological societies exists, as one of EUSEB's first public activities, we invite the secretaries of all such federations to contact Professor Guy Dirheimer, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue René-Descartes, 67084 Strasbourg Cedex, France.

The aims of EUSEB are very much in line with those outlined in Philipson's Commentary, and include in particular the establishment of contacts with national and supranational governmental and private institutions at the highest level, with a view to seeking solutions to the problems confronting basic biological scientists in all disciplines.

HAMISH M. KEIR

(President, EUSEB)

Department of Molecular
and Cell Biology,
University of Aberdeen,
Aberdeen AB9 1AS, UK

Cooling the greenhouse with bioenergy

D. O. Hall, H. E. Mynick and R. H. Williams

Global warming caused by burning fossil fuels could be reduced by the use of biomass for energy. This strategy could be more effective than sequestering carbon by growing more trees.

MOST proposals for reducing global warming have focused on the need to plant more trees in forest reserves, the idea being that carbon dioxide absorption would continue until the trees mature, say for 40–100 years. Although it is recognized that this is not a permanent solution, this 'carbon sequestration' strategy buys time to develop alternative energy sources. Little attention has been paid to another approach to combat global warming, that of substitution of energy derived from biomass for fossil-fuel energy¹. If biomass is grown for energy, with the amount grown equal to that burned for a given period, there would be no net build-up of carbon dioxide in the atmosphere because the amount released in combustion would be compensated for by that absorbed by the biomass during photosynthesis. The potential for reducing carbon dioxide emissions in this way depends on the fuel displaced and on the efficiency with which energy can be produced.

Suppose that the conversion efficiencies are equal. Then a gigajoule (GJ) of biomass substituted for coal would reduce emissions by the carbon content of 1 GJ of coal, about 0.025 tonnes of carbon (tC). Because biomass, with a heating value of 20 GJ per tonne, is 50% carbon, growing 1 GJ of biomass sequesters 0.025 tC. Thus, substituting biomass for coal would be equivalent to carbon sequestration in its effect on atmospheric CO₂. Substituting biomass for petroleum or natural gas would be less effective than carbon sequestration, as these fuels contain less carbon per GJ. Although efficiencies for biomass with commercially available conversion technologies are typically less than for fossil fuels, development of energy-conversion technologies should improve this situation.

Modernization

Although little research and development has been committed to 'modernizing' biomass, there have been serious efforts to 'modernize' coal; much of this technology could be transferred to biomass strategies. The most promising near-term option involves adapting to biomass simplified integrated gasifier/combined-cycle (IGCC) technology (using gas turbines in combined cycles) being developed for power generation with coal². This technology makes it possible to achieve, at the modest scales

needed for biomass power generation, efficiencies higher than those for large coal steam-electric power plants. Biomass versions of this technology should involve lower unit capital cost, allowing it to compete with coal steam-electric power in many circumstances^{3,4}. Biomass versions of simplified IGCC technology could be commercialized more quickly than the corresponding coal versions, because the latter require techniques for sulphur removal that are not yet commercially tested, whereas biomass contains negligible sulphur.

Although they are not competitive at today's low oil prices, synthetic fuels need to be developed to avoid overdependence on oil imports when prices rise once more. The technology for making methanol from biomass is similar to that being developed for coal, as are the conversion efficiencies. Methanol derived via thermochemical processes from biomass, as well as ethanol derived via enzymatic hydrolysis of lignocellulosic feedstocks, could be competitive with gasoline by the year 2000, if the necessary technologies are developed^{5–7}.

Although harvesting, transport and processing requirements make biomass for energy much more costly than growing trees to sequester carbon, energy sales revenues can be taken as a credit against cost. Because the prospects are good that biomass-derived electricity and liquid fuels can be produced competitively, the net cost of offsetting CO₂ emissions by substituting biomass energy for fossil fuels could often be near zero or even negative, and thus lower than the cost of offsetting CO₂ emissions by sequestering carbon in trees¹.

Biomass can play a larger role in reducing global warming when used to displace fossil fuel than when used for sequestration, in part because land can be used indefinitely in displacing fossil CO₂ emissions, whereas CO₂ removal ceases at forest maturity in the 'sequestration' strategy. Also, when biomass is produced for energy markets, producers will seek to maximize the harvestable annual yield rather than the total amount of carbon that can be sequestered in a mature forest. This goal shift will lead them to choose short-rotation woody or herbaceous crops, for which annual yields are 2–3 times as large as for long-rotation species^{8–10}. Furthermore, herbaceous crops can often be

grown at relatively high productivity on crop and pasture lands where the soil and climate conditions are not particularly favourable for growing trees.

Finally, when biomass is used as a fossil-fuel substitute, the land that can be used for biomass production is not restricted to new areas of land for planting 'energy' crops. Agricultural and forest residues that can be economically recovered and low-quality wood resources from existing forests, for example, are potential sources of biomass energy.

In many temperate-zone forests, annual removals of trees are much less than annual growth. Although much of the unharvested stock is too low in quality for traditional forest-products markets, it is well-suited for energy applications, and removal of the low-quality woodstock can simultaneously lead to enhanced yields of high-quality wood¹¹. This increased productivity of high-quality wood in regrowth forests can help ease pressures to exploit original-growth forests, thereby allaying environmental concerns. In practice, the relative contributions to biomass energy — supply of residues, increased production from existing forests and plantation energy crops — will be determined by economics, availability of water and land resources, and constraints posed by environmental concerns.

Potential

The alternative global CO₂ emission projections advanced by the response strategies working group (RSWG) of the Intergovernmental Panel on Climate Change (IPCC)¹² provide a useful context in which to assess the potential for offsetting fossil CO₂ emissions with biomass energy. Global CO₂ emissions could be reduced to half the 1985 level by 2050 by supplementing the response measures of the RSWG "Scenario B" (involving a reversal of deforestation and emphasis of efficient energy use and natural gas) with biomass energy production sufficient to displace by 2050 5.4 Gt per year of carbon from fossil-fuel combustion¹. This might plausibly be achieved by displacing coal with biomass — about one-third of which might come from various agricultural and industrial biomass residues and two-thirds from biomass plantations. Such residues are prime candidates for the first bioenergy systems. The required amount of planta-

tion biomass might be achieved with 600 million hectares of plantations at an average productivity of 12 dry tonnes per ha per year.

This productivity is consistent with what has been achieved only at modest scales. But considering that the era of modern scientific silviculture began only around 1970⁸, and that the growing of herbaceous crops for energy purposes is even more embryonic, this average productivity might plausibly be achieved on a large scale by the middle of the next century. For comparison, average productivities of wheat in the United Kingdom and maize in the United States have more than tripled since the mid-1940s; at present maize yields from 24 million hectares in the United States average 7.5 tonnes per ha per year of grain plus an equal quantity of residues¹. The annual yield of sugar cane, averaged over 17 million hectares of cane harvested globally in 1987, was about 35 dry tonnes per ha per year of total above-ground harvestable plant matter. Moreover, the targeted productivity corresponds to a 0.4% efficiency for converting solar energy into recoverable biomass energy, which is low compared to the practical maximum photosynthetic efficiency under field conditions (5%)¹³ and what has been attained under optimal field conditions (2.4% for Napier grass).

The land area targeted for biomass energy crops in 2050 is large, equivalent to 15 and 40% of the amount of land now in forests and croplands, respectively. Yet estimates of the amount of tropical land potentially available for reforestation are of the order of 800 million hectares¹⁴. Moreover, some of the 1,500 million hectares of tropical grasslands in the world could be used for energy crops such as perennial grasses; at present half this area is burned off each year¹⁵. Considerable land might also be available for energy crops in industrialized countries. In Europe, 15 million hectares of cropland or more would be taken out of production if agricultural surpluses and European Community expenditures on agricultural subsidies were bought under control⁹. In the United States, 30 million hectares of cropland are left idle to reduce production or conserve land; this area could double over the next 25 years. Such considerations suggest that the productivities and land areas associated with this biomass energy scheme are plausible, though ambitious.

The chief uncertainty about the development of biomass for large-scale energy production is whether high productivities can be achieved sustainably over wide areas without damaging the environment. One concern of environmentalists is that biomass energy will reduce biological diversity¹⁶. Certainly,

if monoculture biomass plantations were to replace old-growth natural forests there would be substantial loss of biodiversity. But plantations could be established on deforested or degraded lands, and short-rotation tree crops and various perennial grasses would be an improvement on annual row-crop agriculture.

Experiments in Sweden have shown that in most forests trees grow at rates far below their potential and that nutrient availability is often the most important limiting factor. Under nutrient-optimized conditions all tree species investigated have achieved about the same total biomass yield per unit of light intercepted during the growing season¹⁷. Thus, growing trees under nutrient-optimized conditions could make it possible to achieve high yields with existing species and clones, facilitating the incorporation of pest resistance and other desirable characteristics, and the maintenance of a diverse landscape mosaic. Nutrient-induced yield increases can be achieved without nutrient leaching when good forest management is practiced. But achieving sustainable high yields this way requires matching nutrient applications to the varying need for nutrients^{17,18}.

At present, monocultures are favoured for energy crops, in large part because management techniques are being adapted from monocultural systems for agriculture. Achieving sustainable production and maintaining biological diversity may require polycultural strategies (for example, mixed species with various planting configurations and harvesting methods) for biomass production in many areas.

The nutrient status of afforested lands might be maintained both by recycling nutrients and by choosing suitable mixed species and clones¹⁹. The promise of the latter is suggested by 10-year trials in Hawaii, where yields of 25 dry tonnes per ha per year have been achieved without nitrogen fertilizer when *Eucalyptus* is interplanted with nitrogen-fixing *Albizia* trees²⁰.

Achieving high levels of biological diversity requires maintaining some land in a natural condition. Experience in Sweden suggests that maintaining a modest fraction of forest area in reserves is adequate for this purpose. But research is needed to find out how best to achieve desirable levels of biological diversity under the wide range of conditions where biomass might be grown for energy.

Conclusions

In conclusion, the growing of biomass for energy provided by modern energy-conversion systems would enable biomass to play much wider roles in coping with greenhouse warming than is possible with the growing of trees solely for carbon sequestration. Although carbon-sequestering strategies will be important where the biomass cannot be practically harvested for energy, or where the creation of new forest reserves is deemed desirable for environmental or economic reasons, biomass energy strategies will usually be preferred. Moreover, as biomass energy should often be cheaper than fossil-fuel energy, these strategies will be easier to implement than many other proposed strategies for reducing greenhouse warming.

Bioenergy industries have already been launched in several countries. Nevertheless the techniques and technologies for growing biomass and converting it into modern energy carriers must be more fully developed. If the research and development needed on sustainable production and conversion of biomass is given high priority, and if policies are adopted to nurture the development of bioenergy industries, these industries will be able to innovate and diversify as they grow and mature. □

D. O. Hall is in the Division of Biosphere Sciences, King's College, Campden Hill Road, London W8 7AH, UK. H. E. Mynick and R. H. Williams are at the Center for Energy and Environmental Studies, Princeton University, Princeton, New Jersey 08544, USA.

- Hall, D. O., Mynick, H. E. & Williams, R. H. *Carbon Sequestration vs. Fossil Fuel Subscription: Alternative Roles for Biomass in Coping with Greenhouse Warming*. PU/CEES Report No. 255 (Center for Energy and Environmental Studies, Princeton University, 1990).
- Corman, J. C. *Systems Analysis of Simplified IGCC Plants* (US Department of Energy, Washington, DC, 1986).
- Elliott, P. & Booth, R. *Sustainable Biomass Energy* (Shell International Petroleum, London, 1990).
- Williams, R. H. & Larson, E. D. in *Fuels and Electricity from Renewable Sources of Energy* (eds Johansson, T. B., Kelly, H., Reddy, A. K. N. & Williams, R. H.) (Island, Washington, DC, in the press).
- The Potential of Renewable Energy* (Office of Policy Planning and Analysis, US Department of Energy, Washington, DC, 1990).
- Lynd, L. R., Cushman, J. H., Nichols, R. J. & Wyman, C. W. *Science* **251**, 1318–1323 (1991).
- DeLuchi, M. A., Larson, E. D. & Williams, R. H. PU/CEES Report No. 263 (Center for Energy and Environmental Studies, Princeton University, 1991).
- Kulp, J. L. *The Phytosystem as a Sink for Carbon Dioxide* (Electric Power Research Institute, Palo Alto, 1990).

- Hummel, F. C., Palz, W. & Grassi, G. (eds) *Biomass Forestry in Europe: a Strategy for the Future*, 1–7 (Elsevier Applied Science, London, 1988).
- Cannell, M. G. R. *Scan. J. Forest Res.* **4**, 459–490 (1989).
- Office of Technology Assessment *Energy from Biological Processes* (US Congress, Washington, DC, 1980).
- Intergovernmental Panel on Climate Change *Formulation of Response Strategies, Report Prepared for IPCC by Working Group III* (WMO and UNEP, Geneva, 1990).
- Bolton, J. R. & Hall, D. O. *A. Rev. Energy* **4**, 353–401 (1979).
- Houghton, R. A. *Ambio* **19**, 204–209 (1990).
- Hao, W. M., Liu, M. H. & Crutzen, P. J. in *Proc. Third Int. Symp. Fire Ecology* (ed. Goldammer, J. G.) (Springer, Berlin, in the press).
- Cook, J. H., Beyea, J. & Keeler, K. H. *A. Rev. Energy* (in the press).
- Linder, S. in *Nutrition of Trees* (Marcus Wallenberg Foundation Symposium **6**, 62–87 (Falun, Sweden, 1989)).
- Kimmins, J. P. *Forest Chronical* **66**, 271–280 (1990).
- Mitchell, C. P. *Biomass* **22**, 91–105 (1990).
- Debell, D. S., Whitesell, C. D. & Schubert, T. H. *Forest Sci.* **35**, 64–75 (1989).

Is homosexuality hard-wired?

A suggestion, based on measurements of *post-mortem* brains, that structures in the hypothalamus correlate with sexual behaviour, should be taken seriously. But the meaning is not yet clear.

THE empirical, illogical and unverifiable doctrine that 'every cloud has a silver lining' appears, against the odds, to apply even to the AIDS epidemic. At least, if it were not for the epidemic, Dr Simon LeVay from the Salk Institute at La Jolla, California, would not have been able to examine samples of brain tissue from 16 people known to have been male homosexuals nor to speculate on that basis (*Science* **253**, 1034; 30 August 1991) that male sexual orientation is linked with the size of one of four particular nuclei (or stainably recognizable groups of neurons) in the hypothalamus.

LeVay's claim has naturally attracted a great deal of attention in newspapers around the world. The reasons are not far to seek, for one interpretation of his observations is that genetic males may behave homosexually because of variations in neural structures in the brain that may even be congenital. If that proves indeed to be the case, the tenor of much public argument about the morality of homosexuality would be transformed — or should be. Could Christians, for example, continue to define homosexual behaviour as sinful if it were known to stem from some hard-wired neural structure?

Also naturally, the claim has attracted much scepticism. After all, the numbers are small, the work entails techniques unfamiliar outside neuroanatomy, the outcome of LeVay's study is genuinely surprising — and the implications could be considerable. But, of course, there is no reason why a result that is surprising should be scorned on that account alone, while those who know LeVay attest to the excellence of his work in neuroanatomy. (LeVay also volunteers that he is himself a homosexual.) So scepticism is inappropriate. Getting to grips with the significance of the observations is another matter.

It should be no surprise that there are neural structures that help to determine sexual behaviour. Distinctive neural structures are one obvious means by which gender-specific differences in sexual behaviour could be determined, and the natural place to look for them is in the hypothalamus, with its close links with the pituitary gland, which, with the gonads and to a lesser extent the adrenal glands, is the chief source of hormones regulating the reproductive system.

LeVay has been concerned with four particular nuclei in the hypothalamus (called INH1, 2, 3 and 4) which have

previously been thought to be linked with gender-specific behaviour. He set out to tell whether, instead, they (or at least some of them) are determinants of sexual behaviour. Originally, he says, his ambition had been to compare the sizes of the four nuclei in male and female homosexuals, in each case matched with an appropriate control group. In the event, compiling a set of samples of brain tissue from female homosexuals proved to be impossible (for lesbian behaviour does not transmit human immunodeficiency virus or otherwise declare itself on death). So LeVay's study is based on three sets of brain tissue samples — one from a group of 19 declared male homosexuals; one from a group of 16 males, six of whom died of AIDS and ten of whom died of causes other than AIDS and whose sexual orientation is unknown; and one from a similar group of six women also "presumed heterosexual".

The presumption of heterosexuality in the two control groups sounds fishy, but does not weaken the study. LeVay is out to test the hypothesis that at least one of the four nuclei is, on the average, smaller in homosexual than in heterosexual males. If the 'presumed heterosexual' males included, by chance, some homosexuals, the result would simply be to decrease the significance of whatever difference is found in the comparison.

LeVay makes the technique sound much simpler than it can be. Take a section through the hypothalamus containing the four nuclei, slice than into 3- micrometre sections, stain the nuclear cells and measure the geometrical area projected on each section and then calculate the volume of each nucleus by adding together the areas. The hypothalamus is bilaterally symmetrical, so that each tissue sample should yield two sets of measurements, and LeVay has made bilateral measurements on 70 per cent of his material.

The crucial comparison is between the 16 'presumed heterosexual' and the 19 known homosexual males. LeVay first concludes that there is no significant difference between the sizes of three of the nuclei (INH1, 2 and 4). But there is, he says, a significant difference in the distribution of sizes measured for the third nucleus, INH3, which thus becomes the presumed correlate of sexual orientation in male homosexuals. The claim is based on a one-sided analysis of variance — the sort of thing authors are driven to when

referees will not accept a simple *t* test.

Even so, there will be a great deal of argument about the data. One difficulty is that, in each group, the measured size of INH3 spans a more than twentyfold range, from roughly 0.01 to 0.20 mm³. In other words, male homosexuality is not inconsistent with a large nucleus at the high end of the range, which in itself shows that the third nucleus is not an unambiguous determinant of sexual orientation. The size distributions from the homosexual and heterosexual groups differ chiefly in the clustering of about a third of the measurements around the smaller mean value.

On the assumption that the observations are confirmed (obviously important), what is to be made of LeVay's finding? First, there is the familiar difficulty of telling whether the observed difference is the cause of homosexual behaviour (in some) or the effect thereof. The scatter of the measured sizes suggests that nuclear size, if in any sense a 'cause', is neither a unique nor an unambiguous determinant of homosexual behaviour. Statements in some newspapers that LeVay "has found the seat of homosexuality" simply go too far, even if it is clear that the third nucleus (and probably the other three) are somehow involved with the determination of attributes of sexual behaviour.

There is also the possibility that both the measured differences of size and homosexual orientation are themselves the effects of some other cause. There are suggestive experiments in which it has been shown that male rats castrated at an early age develop female patterns of behaviour, and that females (not only rats) to whom testosterone is administered in early life will develop male behaviour. Plainly, the neural correlates of genetically determined gender are plastic at a sufficiently early stage.

In that light, it is almost certainly too soon to say that LeVay has shown that homosexuality is strictly genetically 'hard-wired', although that may eventually prove to be the case. One distinguished practitioner says of LeVay's work that it puts the nail in the coffin of freudian attempts to account for homosexuality in terms of early interactions between parents and children. But even that is not necessarily true. Plastic structures in the hypothalamus allowing the consequences of early sexual arousal to be made permanent might suit the freudians well.

John Maddox

Initiation without an end

Richard J. Jackson

ON page 90 of this issue¹, Macejak and Sarnow provide definitive evidence of a cellular messenger RNA that undergoes translation not by the well-known ribosome-scanning model but by a mechanism of internal ribosome binding. The result constitutes what is to my knowledge a first for cellular mRNAs; my bet, however, is that there are many more such cases to be discovered.

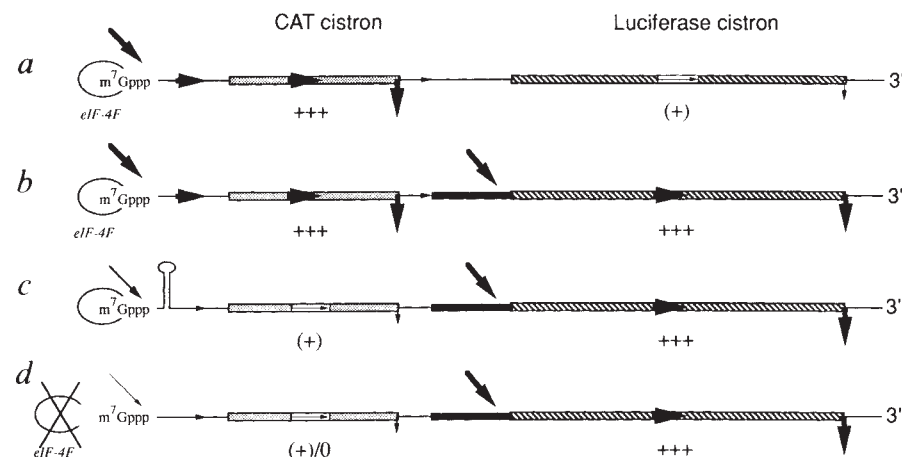
It was once assumed, not without good reason, that eukaryotic ribosomes could gain access to mRNA initiation sites only by interacting first with the capped 5' end of the mRNA, and then scanning linearly in a 5'–3' direction; and that the translation of the second (downstream) cistron of any dicistronic mRNA would be inefficient, because relatively few of the ribosomes which translate the upstream cistron seem to continue scanning through the intercistronic spacer to the downstream cistron². Although there were some claims of exceptions to these rules, few stood the test of close scrutiny until the discovery that initiation on picornavirus RNAs is by a mechanism of internal ribosome entry independent of the 5' end of the RNA (see ref. 3 for review). The fact that no virus-coded products are required for this internal initiation raised the suspicion that some cellular mRNAs might be translated by the same mechanism. Macejak and Sarnow now confirm this conjecture in the case of one mammalian mRNA, that coding for the human immunoglobulin heavy-chain binding protein — BiP, also known as GRP 78, the glucose-regulated protein of relative molecular mass 78,000 (M_r , 78K) — which is thought to play the role of a chaperonin in the endoplasmic reticulum.

In the absence of reliable methods for producing circular mRNAs in high yield in intact cells, most tests of internal initiation involve the study of dicistronic mRNAs (see figure). Macejak and Sarnow transfected mammalian cells with DNA constructs designed to produce such an mRNA with the 5'-proximal cistron coding for chloramphenicol transacetylase (CAT) and the downstream cistron for luciferase. If the 220-nucleotide 5' noncoding region of BiP mRNA was inserted between the two cistrons, the translation of the downstream cistron was selectively stimulated about 30-fold to the extent that luciferase and CAT were produced in approximately equimolar yield (and a similar result was obtained by insertion of part of the 5' noncoding region of poliovirus RNA). An important control

was the demonstration that the mRNAs actually used for luciferase production (those in polyribosomes) were intact and still dicistronic. To show that the high yield of luciferase was not produced by ribosomes scanning through the intercistronic spacer after translation of the

cells, but according to the dicistronic mRNA test this region does not promote internal initiation⁶. In this case it seems that the (lack of) secondary structure in the tripartite leader allows initiation by a mechanism dependent on scanning from the 5' end of the mRNA yet independent of eIF-4F, which is thought to act as an RNA helicase⁷.

Although BiP was the only host-cell protein whose synthesis was seen to continue unabated following poliovirus



Translation of dicistronic messenger RNAs¹: a, control construct with a nonspecific intercistronic sequence; b, with the BiP mRNA 5' noncoding region (shown as a black bar) placed between the two cistrons; c, as b, but with the insertion of a stable hairpin loop near the 5' end; d, as b, but translation in poliovirus-infected cells. Diagonal arrows depict 40S ribosomal subunits binding to the 5' end or to the BiP 5' noncoding sequence (the exact site of entry is unknown); horizontal arrows show 40S subunit scanning through nontranslated regions, or 80S ribosomes translating coding regions; and vertical arrows show the dissociation of both ribosomal subunits at the termination codon. The thickness of the arrows symbolizes the relative number of ribosomes participating in each process. Relative yields of chloramphenicol transacetylase (CAT) and luciferase are scored with the plus symbols.

CAT cistron, a hairpin loop was inserted near the 5' end of the mRNA: as would be expected from previous evidence that such hairpins block ribosome scanning², this insertion greatly reduced the yield of CAT yet did not affect luciferase production (see figure).

The BiP mRNA was singled out as a likely candidate for internal initiation, because the synthesis of a protein of M_r about 78K, astutely identified as BiP, was seen to persist (or even increase) following poliovirus infection⁴. Poliovirus infection of cells (or expression of the poliovirus 2A protease⁵) results in inactivation of initiation factor eIF-4F, the factor responsible for the cap-dependent events in the initiation process. The consequence is a marked reduction in the translation of all host-cell mRNAs (except BiP mRNA) and the CAT cistron of the dicistronic mRNA, but luciferase synthesis is uninhibited¹. However, although persistence of translation of a given mRNA after poliovirus infection is a good indicator of possible internal initiation, it is not an infallible diagnosis: late adenovirus mRNAs with the tripartite leader 5' noncoding region are translatable in poliovirus-infected

infection⁴, this 'screen' would only have detected products of relatively abundant and stable mRNAs. Thus the existence of other mRNAs translated by internal initiation is far from excluded. Indeed, at a recent meeting on translational control held at Sigüenza, Spain, Sarnow reported the intriguing observation that internal initiation can operate efficiently on the long 5' noncoding region of the *Antennapedia* mRNA of *Drosophila*.

Why should certain mRNAs have evolved to be translatable by internal initiation? The most obvious advantage is that it would allow efficient translation under conditions when cap-dependent initiation is inhibited, which can occur not only as a result of eIF-4F degradation (following infection by poliovirus or certain other picornaviruses), but also through dephosphorylation of the smallest subunit of eIF-4F, as happens in mitosis⁸ and in the late stages of infection by adenoviruses⁹. So we might expect to find internal initiation in the case of mRNAs coding for antiviral proteins and for proteins which need to be synthesized efficiently during mitosis, yet neither BiP nor *Antennapedia* clearly fits either of these categories. More exam-

ples of such mRNAs are needed before further speculation is warranted, but in the meantime it is worth noting that internal initiation challenges the usual assumption that polycistronic mRNAs with functional downstream cistrons never occur in eukaryotes.

In the case of picornaviruses, internal initiation requires a length of about 450 nucleotides of the 5' noncoding region. But the sequence requirements are not obvious because the different picornavirus species fall into three groups which share no sequence similarity, apart from a pyrimidine-rich tract some 25 nucleotides upstream of the ribosome entry site, which is at an AUG codon at the 3' extremity of the roughly 450-nucleotide segment³. There is evidence that the secondary structure of the viral 5' noncoding region is essential for internal entry, and it is thought that the important features may be quite short noncontiguous primary sequence motifs held in the appropriate spatial orientation by this higher order structure³. At first sight, the 5' noncoding region of BiP mRNA does not conform to this pattern: it is only 220 nucleotides long (it is not yet known whether the whole length is needed for internal initiation), lacks an appropriately situated pyrimidine-rich tract, and shows no sequence similarity with any picornavirus 5' noncoding region. The 5' noncoding region of BiP mRNA also lacks any AUG codons, whereas the viral sequences have many nonfunctional AUGs.

It will be interesting to see how far internal initiation on the 5' noncoding region of BiP mRNA resembles the picornavirus model. Is the ribosome entry site at the 3' end of the noncoding region? Is secondary structure as important as appears to be the case in the picornaviruses? Does the 5' noncoding region of BiP mRNA bind the same proteins as can be crosslinked to the critical length of about 450 nucleotides of the 5' noncoding regions of different picornaviruses? Can BiP mRNA be translated not only by the internal ribosome entry mechanism but also by scanning from the 5' end, in contrast to the viral RNAs? □

Richard J. Jackson is in the Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK.

And now quasi-semiconductors?

Anders Carlsson

QUASICRYSTALS are a relatively new class of materials that have enough structural order to produce sharp spots in X-ray diffraction patterns, but are not translationally periodic. The effects of this type of structural order on electronic properties such as the electrical conductivity are only just beginning to be

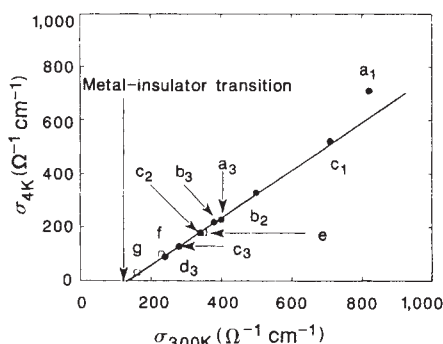


FIG. 1 Comparison of room- and low-temperature conductivity for quasicrystals approaching a metal-insulator transition. Samples a–d are AlCuFe quasicrystals with varying compositions; annealing treatments are specified by numbers. Samples e–g are AlCuRu quasicrystals.

understood. T. Klein, C. Berger, D. Mayou and F. Cyrot-Lackmann, of the Centre National de la Recherche Scientifique in Grenoble, demonstrate that the electronic behaviour of some nearly-perfect icosahedral quasicrystals is rapidly approaching the borderline between metals and semiconductors (*Phys. Rev. Lett.* **66**, 2907–2910; 1991). This news is remarkable, because there are no known semiconducting or insulating quasicrystals. It suggests that further investigation of high-quality quasicrystals may lead to quite exotic electronic properties.

The Grenoble group measured the electrical conductivity, and other properties related to the conduction electrons, of several quasicrystals with composition close to $\text{Al}_{63}\text{Cu}_{25}\text{Fe}_{12}$. The samples were obtained using a variety of annealing treatments to reduce the number of defects (see the News and Views by D. P. Di Vincenzo *Nature* **340**, 504–505; 1989). The lowest conductivities were obtained for the composition $\text{Al}_{62.5}\text{Cu}_{25}\text{Fe}_{12.5}$, after annealing for a few hours at 800 °C. The value of the conductivity below 1 K for this sample was approximately $100 \Omega^{-1}\text{cm}^{-1}$, which is extremely small in comparison with typical metals (see box, over). Similar results had previously been obtained for AlCuRu quasicrystals by B. Biggs, S. Poon and N. Munirathnam (*Phys. Rev. Lett.* **65**, 2700–2703; 1990).

A vital clue to understanding the re-

sults is that in both classes of quasicrystals the conductivity is extremely sensitive to composition. For example, in AlCuFe, varying the composition by only half a per cent to form $\text{Al}_{63}\text{Cu}_{25}\text{Fe}_{12}$ doubles the low-temperature conductivity. This sensitivity to composition is reminiscent of doping effects in semiconductors, in which small impurity concentrations can cause enormous changes in conductivity. Also, either lowering the temperature or improving the sample quality by annealing reduces the conductivity. These features are at variance with the behaviour of typical metals, in which improved sample quality and lowered temperature reduce the random ion-electron scattering in the sample and thus raise the conductivity.

Combined analysis of the low-temperature (4 K) and room-temperature results for AlCuFe and AlCuRu leads Klein *et al.* to the conclusion that a metal-insulator transition is close at hand. This analysis is based on the idea that the factors (other than thermal disorder) that reduce the conductivity at room temperature should do so even more at lower temperature. For example, small densities of conduction-electrons become reduced further at low temperature because of the absence of thermal excitation. Also, theoretical analysis ('weak-localization' theory) indicates that scattering due to inherent (not thermal) disorder in the material can be enhanced at low temperatures, because of quantum-mechanical coherency between scattering events. Thus the high- and low-temperature conductivities should be strongly correlated with each other. In fact, theoretical calculations, as well as empirical correlations, have suggested that the zero-temperature conductivity should vanish for a characteristic finite value of the room-temperature conductivity. This would result in a metal-insulator transition. To explore the possibility of such a transition, Klein *et al.* prepared a plot which shows a strong correlation between the room-temperature and 4-K conductivity (Fig. 1). The plot suggests that bringing the room-temperature conductivity down by another 20 per cent from the lowest observed value would make the zero-temperature conductivity zero, and the material would be the first quasicrystalline semiconductor.

These results cap a trend of observations on the electronic properties of quasicrystals made by several groups over the past five years. The first were made on samples poor in quality in comparison with what can be achieved

1. Macejak, D. G. & Sarnow, P. *Nature* **353**, 90–94 (1991).
2. Kozak, M. J. *Cell Biol.* **108**, 229–241 (1989).
3. Jackson, R. J., Howell, M. T. & Kaminski, A. *Trends biochem. Sci.* **15**, 477–483 (1990).
4. Sarnow, P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5795–5799 (1989).
5. Sun, X.-H. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2143–2146 (1989).
6. Dolph, P. J., Huang, J. & Schneider, R. J. *J. Virol.* **64**, 2669–2677 (1990).
7. Rozen, F. *et al. Molec. cell. Biol.* **10**, 1134–1144 (1990).
8. Bonneau, A. M. & Sonenberg, N. *J. biol. Chem.* **262**, 11139 (1987).
9. Huang, J. & Schneider, R. J. *Cell* **65**, 287–280 (1991).

today; their conductivities were typically in the range $5,000\text{--}20,000\ \Omega^{-1}\text{cm}^{-1}$, much greater than those in the new high-quality samples. Higher conductivities are usually associated with quasicrystals containing only simple metals (such as AlMgZn), not transition metals. Where comparison is possible, poorer-quality quasicrystals generally have electronic properties similar to those of metallic glasses of the same composition. Even with simple-metal quasicrystals, increased structural perfection reduces conductivity. But the effects are more pronounced in quasicrystals that contain transition metals.

The electronic properties of 'rational approximants' have also been measured recently. These are large-unit-cell crystals whose local geometry resembles that of the quasicrystals. Their electrical conductivities are, in general, close to those of the high-quality quasicrystals. Thus one may roughly summarize the experimental data by saying the electronic properties of stable high-quality quasicrystals resemble those of related crystals; those of poor-quality quasicrystals resemble those of related glasses.

Although we are far from making a quantitative theoretical prediction of the electrical conductivity of a transition-metal-containing quasicrystal, some of the underlying physical mechanisms are understood. The best-studied involves structurally induced peaks and valleys in the energy distribution of the quantum-mechanical electron levels (or density-of-states (DOS); see Fig. 2). Scattering of electrons by various Fourier components of the potential they see due to the ions leads to oscillations at characteristic energies which are determined by the

density of atoms and their geometrical arrangement. The highest occupied electron states, whose energy is known as the Fermi level, dominate the conduction processes. If the Fermi level lies in a valley, or quasigap, in the DOS, then the conductivity is reduced.

A well-known crystalline example of the effect is bismuth, which has a rhombohedral structure, with two atoms per unit cell. This structure is almost simple cubic except for a small distortion. If bismuth were simple cubic, it would have a high value of the Fermi-level electron DOS, corresponding to a typical metallic conductivity. But the distortion causes a dramatic reduction of the DOS by allowing additional scattering process. Consequently the conductivity is reduced to a semi-metallic value.

According to the scattering theory, changing the composition of the quasicrystal should empty or fill electron states, shifting the Fermi level and thus changing the conductivity rapidly. Such effects are confirmed experimentally by the new AlCuFe and AlCuRu results, and in several other cases. Furthermore, scattering theory suggests that strong-scattering ions, such as transition metals, should produce stronger DOS oscillations than the weak-scattering simple metals. Again, this prediction is confirmed by the low conductivities seen in AlCuFe and AlCuRu.

Additional support for the scattering theory comes from precise numerical calculations of the DOS in the rational-approximant structures. In all of the stable examples studied, very pronounced oscillations are seen. The Fermi level lies in a quasigap. In fact, the theory suggests that the stability of the

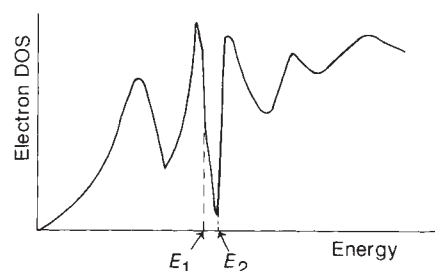


FIG. 2 Peaks and valleys in electronic density of states (DOS) distribution. If the Fermi level lies at E_1 , high conductivity is obtained. At E_2 it results in low conductivity.

rational-approximant phases is partly attributable to the quasigap. This effect is analogous to the particular stability of covalent materials, in which there is actually a gap between the occupied and empty electron levels. The correlation between stability and the Fermi-level quasigap can explain why the stable quasicrystals tend to have the lowest conductivities, because of the reduced value of the Fermi-level DOS.

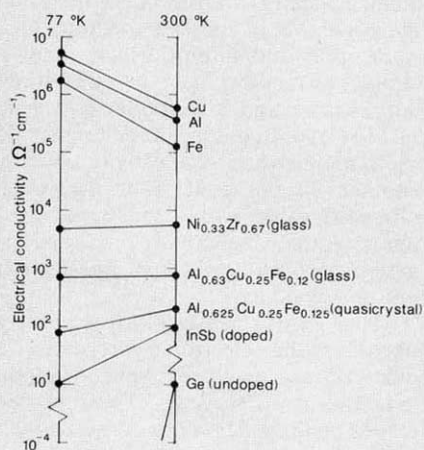
The high precision of the measurements challenges theorists to explain the near-semiconducting behaviour in a more quantitative fashion. The strong similarity between the quasicrystal properties and those of related rational-approximant crystals suggests that the latter will be an essential intermediate stage in elucidating the effects of quasicrystallinity. Existing techniques for crystal-structure determination are sufficiently accurate that atomic positions in the crystals, unlike quasicrystals, can be precisely determined. Furthermore, theoretical techniques for calculating electronic properties of crystals are much more reliable than those for quasicrystals. With the crystalline results as a stepping-stone, we may begin to be able to predict the differences between the quasicrystalline phase and crystalline phases.

The experimental results discussed here show that quasicrystalline ordering can have spectacular effects on the electronic structure of a material. They are the starting point in a search for semi-conducting quasicrystals, whose conductivity would vanish as absolute zero is approached. Materials at the borderline between metallic and semiconducting behaviour can display other fascinating properties as well. For example, some transition-metal oxides that are borderline metals at room temperature become high-temperature superconductors at liquid-nitrogen temperatures. It may be that further study of highly perfect quasicrystals will reveal equally startling behaviour. □

Anders Carlsson is in the Department of Physics, Washington University, St Louis, Missouri 63130-4899, USA.

Temperature dependence of electrical conductivity

THE three constituents of the AlCuFe quasicrystal have much higher conductivities than the quasicrystal itself. In the elemental constituents, the resistivity is caused by scattering due to the random thermal displacements of the ions, and thus the conductivity increases as the



temperature is lowered. In glasses formed by two or more metallic constituents, the conductivity is much poorer, because of scattering due to the disorder inherent in the glassy structure. Also, the temperature dependence of the conductivity is much reduced, and can even change sign, as indicated by the NiZr example. The AlCuFe glass shown has a behaviour generally in line with known observations on metallic glasses. In the quasicrystal, however, the conductivity is even lower than that of the glass, and the conductivity drops markedly with decreasing temperature. The resistivity is, in fact, approaching typical values for small-gap semiconductors, such as InSb. The Ge point illustrates the behaviour of larger-gap semiconductors, which have very low conductivities (in the absence of dopant impurities). Their conductivity drops off very rapidly with decreasing temperature because of the exponentially decreasing number of free charge carriers.

A case of missing marsupials

Jared M. Diamond

AUSTRALIAN marsupials are famous for their convergence on placental mammals, most closely in the marsupial 'mole', 'mice' and 'cats', and less closely in the convergence of kangaroos on deer. These equivalents make it all the more surprising that Australia lacks marsupial tigers or other big native carnivores. Might such creatures have evolved, only to become extinct before the first European explorers landed? By bringing together knowledge of Australia's Pleistocene fossil carnivores, Flannery¹ has illustrated how analysis of extinction and species diversity on short timescales by community ecologists can be extended to evolutionary and geological times.

To understand why such trophic imbalances in a fauna are theoretically interesting, consider Brown's worldwide analysis² of the extant large-mammal distributions tabulated by Rusler. The percentage of a continent's mammal genera that exceeds 11 kg in mass increases with continental area, from 0 for the 'mini-continents' of Madagascar and New Guinea, to 3 for Australia, about 10 for North and South America, and 20 for the two largest continents of Africa and Eurasia. In explanation, recall that a population's risk of extinction over a few years declines steeply with population size, for the obvious reason that two individuals are more likely to happen to die simultaneously than are 10. But populations of large animals inevitably consist on the average of fewer individuals than populations of small animals. Thus isolated populations of large animals are disproportionately at risk of extinction in national parks and on land-bridge islands, on timescales ranging from decades to thousands of years³. Brown's analysis suggested that this proneness of large mammals to become extinct also applies on timescales of millions of years.

By the same reasoning, the ecological pyramid guarantees that carnivores will generally be rarer than herbivores of similar size. One would therefore expect large carnivores to be especially under-represented on small land masses. Flannery's synthesis let him test this expectation for Pleistocene Australia (joined at that time to Tasmania and New Guinea).

Australia's sole modern mammalian carnivore exceeding 11 kg was the recently extinct thylacine or marsupial

'wolf'. Even relaxing the size limit to 5 kg adds only two more species, the extant Tasmanian devil and cat-like spotted-tailed quoll. Flannery shows that this modern situation is not an artefact of Pleistocene decimations of large mammalian carnivores, because Pleistocene Australia had only one or two other such species, albeit remarkable ones.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Cold bloodedness and a low metabolic rate may explain the isolated evolution of the Komodo dragon.

One was the so-called marsupial 'lion' (really more like a leopard), with large slicing premolars and clawed paws. The other was an enigmatic, possibly carnivorous, giant kangaroo whose teeth suggest that it may have eaten small vertebrates as well as the plant matter beloved of normal kangaroos. The net result is that Australia had only one-fifth as many species of large mammalian carnivores as do East Africa, North America or Thailand today. The same conclusion of Australian impoverishment is reached if one instead calculates the ratio of carnivores to herbivores among large mammal species — it is 1:4 for East Africa, but only 1:15 for Pleistocene Australia.

Dependence

All this fits in well with the view that the steep dependence of extinction probability on population size operates in evolutionary as well as ecological time. Big carnivorous animals were just too likely to become extinct on the smallest continent to evolve or persist there. Flannery, however, points out two additional factors. First, Australian environments are

notorious for their low productivity stemming from low rainfall and nutrient-poor soils, so Australian population sizes would be even lower than expected from Australia's small area alone (the effects of productivity on species diversity have been well-documented for modern ecological communities⁴). Second, Australia is also notorious for its year-to-year fluctuations in rainfall associated with El Niño events. One expects theoretically, and observes empirically, that populations whose numbers undergo wide temporal fluctuations (for example, as a consequence of fluctuations in primary productivity) are more liable to extinction than are stable populations⁵.

Gerald Cubitt/Bruce Coleman Ltd

Anyone who has visited Australia will realize, though, that an analysis focusing on mammals alone would be incomplete. Australia is distinguished today not only by a deficit of large carnivorous mammals, but also by an excess of large carnivorous reptiles, including 10 species each of pythons and varanid lizards weighing over 5 kg. Pleistocene Australia was even more spectacular in this regard, being endowed with a very large python (over 50 kg), an even larger land crocodile, and a gigantic varanid lizard like an oversized version of the world's largest extant lizard, the Komodo dragon (see figure).

But because cold-blooded animals have lower metabolic rates than warm-blooded ones of the same size, they can survive longer periods of starvation and maintain higher numbers relative to their prey. So it makes sense that the smallest and least productive continent, and the one subject to the greatest fluctuations in rainfall, is enriched in large cold-blooded carnivores. This fact may also explain the puzzling evolution of the Komodo dragon on a small island (Indonesia's Flores group) rather than on a continent, and may also account for a large fossil relative of the dragon recently discovered on New Caledonia.

In short, Flannery's study now invites palaeontologists as well as ecologists to examine systematically how trophic compositions, size distributions and ectotherm: endotherm ratios of faunas vary among islands and continents. □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024, USA.

1. Flannery, T. *Australian Nat. Hist.* **23**, 722–729 (1991).
2. Brown, J. H. *Biol. J. Linn. Soc.* **28**, 231–251 (1986).
3. Diamond, J. M. in *Extinctions* (ed. Nitecki, M. H.) 191–246 (Univ. Chicago Press, 1984).
4. Wright, D. H. *Oikos* **41**, 496–506 (1983).
5. Pimm, S. L. et al. *Am. Nat.* **132**, 757–785 (1988).

A sideways look at quasars

Meg Urry

THE idea that the several different kinds of active galactic nucleus may actually be the same astronomical phenomenon viewed at different orientations follows naturally from the fact that many have roughly cylindrical symmetry. Various schemes for relating the different types ('unified theories') have been suggested over the past decade. Although some are likely to be right, the current problem is finding the direct evidence to show which. Recent studies¹⁻⁵ of the nearby powerful radio galaxy, Cygnus A, constitute a plausible, but still circumstantial argument that it is a quasar viewed from the side.

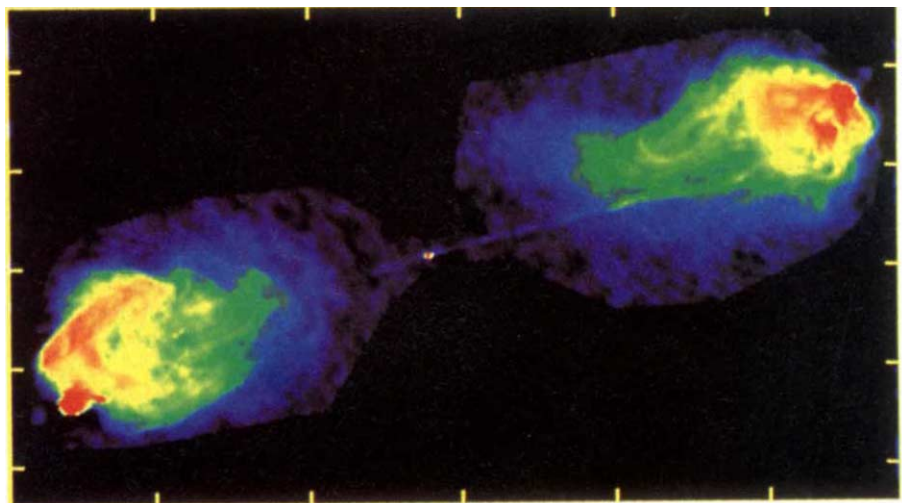
The one thing all active galaxies have in common is a source of exceptional brilliance concentrated in the nucleus of each. But beyond that unity, there is a plethora of varieties, and the question is which variations are fundamental and which depend on the way we see things. The impetus for a particular unified theory usually starts with the identification of similarity between roughly isotropic properties of different objects. For example, large-scale features such as extended regions of radio power or of gas radiating narrow-line spectra or even the host galaxies may resemble one another. Observations at some wavelengths are relatively unaffected by absorption (radio, far infrared and hard X-rays) and so allow an unobscured (and so orientation-independent) view of the active nucleus.

Some of the proposed unification schemes are clearly viable in the sense that the isotropic properties of each class match and their number densities and luminosities are as expected. What is needed now is direct evidence indicating that a particular radio galaxy, say, actually harbours a quasar or BL Lacertae object at its centre. For example, the two types of Seyfert galaxy may be fundamentally the same, as indicated by optical spectropolarimetry⁶ of Seyfert 2s. The results imply that they contain a broad-line region (like Seyfert 1s) that is hidden by obscuring material, and so visible only in scattered (and hence polarized) light.

Radio-loud active galaxies may be affected in a similar way by obscuration, as well as by relativistic beaming along the jets that emanate from their nuclei. There is increasing support for the hypothesis that BL Lac objects (a class of lineless, rapidly variable, highly polarized radio-loud active galactic nucleus) are low-luminosity radio galaxies viewed along their jet axes^{7,8}. Recent narrow-band imaging and spectroscopy of the

optical filaments surrounding the low-luminosity radio galaxy Centaurus A suggest that the photoionizing source is an intense, highly anisotropic cone of light directed towards the filaments and away from us⁹, just as expected in this unified scheme.

At higher luminosities, radio-loud quasars (like BL Lacs, flat-spectrum radio quasars are generally polarized and variable) may be aligned high-luminosity radio galaxies^{10,11}. Many high-luminosity



A radio map of Cygnus A, showing the giant radio lobes and bright nucleus.

radio galaxies, particularly those at high redshift (great distances), show optical and infrared light that is generally aligned with the radio axis¹². In several cases (including³ Cygnus A), imaging polarimetry of the extended optical continuum suggests that it consists at least partly of scattered light from an obscured nucleus, although, probably, jet-induced star formation also contributes^{3,12}. At infrared wavelengths, obscuration is orders of magnitude less than at visible wavelengths, so it may be possible to detect the hidden nuclear continuum source and the broad-line region through directly transmitted light. At X-ray energies, an active nucleus can be detected through its compact, variable, non-thermal emission.

All of these tests have been applied to Cygnus A. As the nearest example of a high-luminosity radio galaxy, Cygnus A is the most luminous radio galaxy in the local Universe. Its disturbed optical morphology, powerful double radio lobes (see figure), and strong narrow-line emission are similar to those of high-redshift radio galaxies¹², but because of its proximity it can be explored with much higher spatial resolution and better signal-to-noise. If powerful radio galaxies contain hidden quasars, Cygnus A

should be the easiest case to prove.

So what evidence has been garnered? The extended featureless optical continuum led to the first suggestion¹ that Cygnus A has an obscured quasar nucleus, and this picture is supported by imaging polarimetry³ (although the spectropolarimetric results have been interpreted differently¹³). Cygnus A also appears to have a nonthermal, presumably nuclear, X-ray source², as there is an excess of hard X-rays above the background thermal emission that dominates that region of the sky; observations of X-ray variability would confirm this.

Now infrared observations are also

being interpreted as giving evidence of a hidden nucleus. Djorgovski *et al.*⁴ report the discovery of an unresolved, point source of infrared light coincident with the radio core in Cygnus A. Comparing the measured infrared luminosity with that expected from an extrapolation of a quasar-like radio spectrum, the authors argue that the point source is strongly obscured (by 20 orders of magnitude) by material along our line of sight. Ward *et al.*⁵ apply similar indirect arguments to infrared spectroscopy of the Paschen series of hydrogen lines⁵. If Cygnus A obeys the same correlation between X-ray flux and H α as do other active galactic nuclei, and the intrinsic ratio of Pa α to H α flux is normal, then the absence of broad Pa α indicates at least about 10 orders of magnitude of visual extinction toward the putative broad-line region. Similarly, if the correlation between X-ray and infrared fluxes for radio-loud quasars applies to Cygnus A, there must be about 20 orders of magnitude of visual extinction toward the continuum source. If the obscuration is not considerably greater, the broad-line region should be directly detectable with a modest increase in sensitivity in the infrared.

So far, the picture of a buried quasar

NATURE · VOL 353 · 5 SEPTEMBER 1991

OBITUARY

Allan Charles Wilson (1934–1991)

in Cygnus A is a consistent one. For a normal gas-to-dust ratio, the proposed continuum extinction is roughly consistent with the observed absorbing column density along the line of sight to the X-ray source² ($8 \times 10^{22} \text{ cm}^{-2}$). Furthermore, there is an excess of far-infrared luminosity that is comparable to the amount of shorter-wave radiation that must be absorbed, as would be expected if it were re-radiated by dust⁵. (Interestingly, the far-infrared peak is relatively narrow, indicating that the dust has a well-defined temperature, of about 75 K, so that its spatial distribution is probably narrow, like a torus, rather than extended, as a warped disk would be.)

Although these observations offer a tantalizing glimpse into the possibilities of verifying unified schemes for active galactic nuclei, additional questions remain. Foremost, can the broad-line region that is characteristic of all quasars be detected directly in Cygnus A? Transmitted Paschen lines might be seen or the spectrum of scattered, polarized light could be searched for broad lines. If unification survives that test, the relation of Cygnus A to the much more numerous high-redshift radio galaxies (and hence high-redshift quasars) is still not understood. And even if the global unified scheme for radio-galaxies and quasar proves true, a detailed understanding of the structure of radio galaxies remains an important goal. Why are many intrinsically asymmetric? What is the geometry of the obscuring material and nuclear continuum source? And is the continuum source itself intrinsically anisotropic? How do radio galaxies form and evolve? None of these questions will be easy to answer, but to start by demonstrating the unification between radio galaxies and quasars in the case of Cygnus A would be an enormous step forward. □

Meg Urry is at the Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA.

ALLAN Wilson, the most influential figure in the empirical study of molecular evolution, died on 21 July from complications resulting from bone-marrow transplantation for treatment of leukaemia. He was at the height of his productivity.

Born in Ngaruawahia, New Zealand, he remained a New Zealand citizen. He received his PhD in 1961 from the University of California at Berkeley, under Arthur Pardee. After postdoctoral study with Nathan Kaplan at Brandeis University, he joined his old department at Berkeley, where he remained, building up the most prominent of the groups working on molecular evolution.

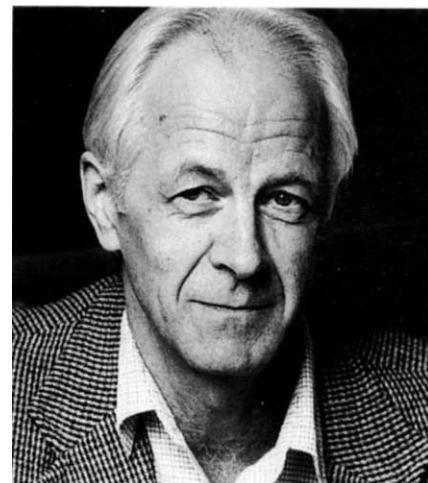
Systematists, microbiologists and biochemists came to Berkeley to learn the techniques involved. But Wilson's influence went beyond this. While others concentrated on what evolution could tell them about molecules, Allan always looked for ways that molecules could say something about evolution.

His greatest effect was on the study of human evolution. In early work with Vincent Sarich, he used immunological methods to study hominid proteins and come to the conclusion that humans and African apes shared a common ancestor only five million years ago. To physical anthropologists this date was far too recent; it has gained increasing acceptance. With Mary-Claire King he used electrophoresis to estimate that human and chimpanzee proteins differ by only 1 per cent, a remarkably small figure. And with Wes Brown and Stephen Ferris, he made a restriction sites phylogeny that grouped humans with the African apes.

The disputes over these results paled in comparison with the uproar that attended appearance of the paper by Cann, Stoneking and Wilson (*Nature* 325, 31–36; 1987), in which restriction-site variation in human mitochondrial lineages was employed to place the common mitochondrial ancestor in Africa 200,000 years ago. Much of the publicity was based on the mistaken belief that this mitochondrial 'Eve' was the only female of that generation who had modern descendants, thereby resonating with popular notions of Earth-mother goddesses and creationism. The work actually implied that population sizes were in the low thousands at that time. A more controversial implication is that *Homo erectus* and Neanderthals in places other than Africa contributed few or no mitochondria to present-day humans. There is no standard error on the dates, so that we do not yet know how strong this conclusion is.

Wilson's research, much of it carried out in collaboration with his associate Ellen Prager, proceeded from immunological methods to restriction sites to DNA

sequencing. His laboratory was one of the first to use the polymerase chain reaction, and had a close involvement in simplifying mathematical methods for analysing data. The laboratory was a leader in sequencing 'fossil DNA', with studies on the extinct quagga by Russell Higuchi and human mummies by Svante Pääbo that both made news and broke new ground scientifically.



Allan Wilson — 'father' of 'Eve'.

Much of the work depended on the 'molecular clock', for which Wilson was one of the principal advocates. It is clear that this 'clock' runs at different rates in different groups; it is not universal. Morphologists tend to dismiss it as simply wrong, but for related species it is a useful and fruitful approximation. Wilson spent considerable effort in trying to quantify morphological evolution. He argued, for example, that frogs had diversified less than mammals in the same time, and speculated that this was because morphology interacted with mammalian social behaviour. He also argued that morphological change is a consequence of a specialized class of regulatory mutations.

Allan Wilson was a retiring man who yet managed to publicize his work to great effect; he was simultaneously daring in his hypotheses and sensitive to criticism. He and his wife of 31 years, Leona, faced his final struggle together with great courage. As he prepared to undergo transplantation, he became interested in HLA and expressed the view that African mitochondrial diversity implies that there should be many more HLA alleles waiting to be discovered there. It is typical of Allan that he leaves a legacy of challenging hypotheses that encourage us to collect more and better data.

Joseph Felsenstein

Joseph Felsenstein is in the Department of Genetics, University of Washington, Seattle, Washington 98195, USA.

- Pierce, M. J. & Stockton, A. *Astrophys. J.* **305**, 204–213 (1986).
- Arnaud, K. A. et al. *Mon. Not. R. astr. Soc.* **227**, 241–256 (1987).
- Tadhunter, C. N., Scarrott, S. M. & Rolph, C. D. *Mon. Not. R. astr. Soc.* **246**, 163–168 (1990).
- Djorgovski, S., Weir, N., Matthews, K. & Graham, J. R. *Astrophys. J.* **372**, L67–L70 (1991).
- Ward, M. J., Blanco, P. R., Wilson, A. S. & Nishida, M. *Astrophys. J.* (in the press).
- Antonucci, R. R. J. & Miller, J. S. *Astrophys. J.* **297**, 621–633 (1985).
- Browne, I. W. A. in *BL Lacs Objects* (eds Maraschi, L. et al.) 401–411 (Springer, New York, 1989).
- Urry, C. M., Padovani, P. & Stickel, M. *Astrophys. J.* (in the press).
- Morganti, R. et al. *Mon. Not. R. astr. Soc.* **249**, 91–112 (1991).
- Peacock, J. A. in *Astrophysical Jets and their Engines* (ed. Kundt, W.) 185 (Reidel, Dordrecht, 1987).
- Barthel, P. D. *Astrophys. J.* **336**, 606–613 (1989).
- Chambers, K. C., Miley, G. K. & van Bruegel, W. *Nature* **329**, 604–611 (1987).
- Goodrich, R. W. & Miller, J. S. *Astrophys. J.* **346**, L21–L24 (1989).

A portable palette

THE first issue of *The Plant Journal* comes adorned with a multi-coloured display of mutant flowers prepared by Rosemary Carpenter and colleagues of the John Innes Institute in Norwich (1, 59–69; 1991). The new pigmentation mutants are generated by inserting transposons at key sites in the genes of *Antirrhinum majus*. Transposons

are mobile genetic elements that can insert themselves at random in the genome, inactivating genes whose coding sequence is disrupted by the insertion or activating others when a promoter or transcriptional activator fortuitously results from the mutation. The phenolic pigment anthocyanin is responsible for the colour of *Antirrhinum* flowers, and strategic insertion of the transposons Tam1,2,3,4 into genes encoding enzymes active in anthocyanin synthesis gives the



variety of flowers shown.

As the red wild-type flower (top, left) develops, anthocyanin synthesis normally follows a temporal and spatial pattern which defines different areas of the flower corolla. Synthesis starts in a ring at the base of the tube and in the lobes; the tube then realizes its full colouring. The other flowers result from transposon mutagenesis in different loci. The *nivea* locus encoding a synthase for the pigment precursor chalcone gives an albino

bloom (top, centre) when it carries Tam4 in its first exon: the mutant chalcone synthase is incapacitated by being longer than it should be. A Tam1 insert in *incol-orata* gives tinged ivory flowers (top, right) which result from a later block in the pigment synthesis pathway. Disruption by Tam2 of *delila*, a gene that regulates the spatial

distribution of anthocyanin, results in a bloom with coloured lobes and an ivory tube (bottom, left). The gene affected in the last two flowers, *daphne*, participates in the synthesis of flavones, the co-pigments that add brilliance to the anthocyanin colour. Here the flavones are missing, hence the dullness of the flowers. With the genetic engineer's increasing dexterity in manipulating pigmentation, who says the leopard still cannot change its spots? □

CELL BIOLOGY

The Janus factor

R. M. Warn

SCATTER factor and hepatocyte growth factor are two proteins with very different origins from which it would have been rather hard to deduce a common link. However last year Gherardi and Stoker¹ proposed, on the basis of sequence similarity of the N-termini, that the two were either similar molecules or even identical. The second proposal is now confirmed by Weidner *et al.*² who demonstrate, by a variety of criteria, that they are the same protein, and also give a chromosomal location — region 11.2–21 on the long arm of chromosome 7 — for the human factor.

Scatter factor was originally identified as a protein secreted by fibroblasts which disperses epithelial colonies of a variety of cell types in culture³. Subsequently it was found that smooth muscle cells also secrete the factor and endothelial cells too are scattered by it⁴. Scattered cells are highly motile, continuously extending processes by mechanisms reminiscent of neuron growth cones and retracting them in similar manner to the snap-back of fibroblast processes⁵. But scatter factor was not found to act as a mitogen for at least one kidney epithelial line susceptible to colony dispersion⁶.

In contrast, hepatocyte growth factor was first recognized as a constituent of the serum of patients with fulminant

liver failure⁷ and that of partially hepatectomized rats⁸, and was further found to be a potent mitogen of hepatocytes in culture. It was therefore initially recognized as a growth factor involved in liver regeneration. No effects on cell motility were reported for the assay systems used. The two molecules were only realized to be the same when sequence data became available. The final demonstration is that fibroblast scatter factor stimulates primary hepatocytes to divide and recombinant hepatocyte growth factor causes dispersion *in vitro* of a variety of epithelial colonies^{2,9}.

The molecule, which perhaps should now be known as HGF-SF, is a heterodimer; purified human factor shows several forms and consists of a large and a small subunit with relative molecular masses of 54,000–65,000 and 31,000–35,000 respectively. The large subunit contains four kringle modules, domains that are present in a variety of proteins including plasminogen and plasminogen activator and that are implicated in protein–protein interactions¹⁰. The smaller subunit resembles the catalytic subunit of plasmin and other serine proteases, with the difference that the histidine and serine residues of the active site are replaced and so the molecule shows no protease activity.

Only a few months ago, the receptor for HGF-SF was identified by two groups^{11,12}, and it turns out that the receptor is the product of the *c-met* proto-oncogene; *c-met* has been a receptor in search of a ligand for some time, having been first identified in 1984 (ref. 13). It is a membrane-bound tyrosine kinase of relative molecular mass 190,000, and thus has significant structural and functional features in common with the receptors of other growth factors, including platelet-derived and epidermal growth factors.

Now that HGF-SF is known to be a single entity, what conclusions can be drawn concerning the dual nature of the molecule? That certain growth factors promote cell motility has been known for some time (see refs 9 and 14 for reviews). But it is remarkable that HGF-SF can stimulate the motility of some epithelial cells without having any effect on mitosis. What kinds of signal cascade are activated by *c-met* and how they can be varied to have differential effects on cell division and motility are now questions of some interest.

Finally, what is the potential significance of the regulation of cell division and motility by HGF-SF in tumorigenesis, where both cell growth and movement are frequently grossly altered? The molecule induces the migration of transformed and non-transformed epithelial cells into collagen matrices, mimicking the effects of tumour cells *in vitro*¹⁵. Thus HGF-SF may initiate or enhance the metastatic

spread of a tumour simply as the result of its presence in the local environment. Furthermore, a tumour might become responsive to the presence of the factor as the result of the expression or amplification of *c-met* receptors present on the cell surface. As a consequence an invasive phenotype would be induced. A further possibility is that the tumour itself might come to synthesize the factor. Up to now HGF-SF has been found

largely to be synthesized by one cell type but to have its effects on another (that is, it is paracrine in action). But one report has demonstrated the autocrine synthesis of the cytokine by a keratinocyte line¹⁶ and a similar mechanism may operate in the development of some tumours. □

R. M. Warn is in the School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

RÉSUMÉ

In a nutshell

A TECTONIC history of the Himalayas based on a single mineral grain is the bold promise of F. M. Richter *et al.* (*Earth planet. Sci. Lett.* **105**, 266–278; 1991). The authors' claim is that different domains in their feldspar sample retain ⁴⁰Ar, a radiogenic daughter of ⁴⁰K, up to different maximum temperatures. By step heating the sample and analysing the composition of gas drawn off it at each temperature, they reconstruct the (reverse) cooling history of the grain based on the range of argon isotope ages. The authors then deduce that up to around 20 million years ago, the sample was buried 10 km below the Earth's surface; 5 million years later it was only 2 km deep. The simple inference is that this was the period over which rapid tectonic uplift occurred, accompanied by erosion of the overlying rocks from the evolving mountains.

Counter current

A PROMISING form of treatment to halt the lung disease in cystic fibrosis (CF) patients is described by M. R. Knowles *et al.* in the *New England Journal of Medicine* (**325**, 533–538; 1991). The severe bacterial lung infections associated with CF are closely tied to defective chloride secretion and raised sodium absorption through the airway epithelia, which alter the chemical environment in the lung. In preliminary tests on cultured nasal epithelial cells from 12 CF patients, Knowles *et al.* find that a combination of the nucleotide adenosine (or uridine) triphosphate with amiloride induces chloride secretion. The extracellular nucleotides possibly bypass the defective CF chloride channel by stimulating a class of cell-surface nucleotide receptor. The results raise hope that an aerosol preparation may be an effective form of therapy.

Outer limit

THE discovery of a quasar at record-breaking distance at the start of a large-scale quasar survey bodes well for the investigators, D. P. Schneider, M. Schmidt and J. E. Gunn (*Astr. J.* **102**, 837–840; 1991). The survey is intended to discover quasars at redshifts of 4.0–5.5 — at the limits of the observable Universe. Until the new discovery, the most distant known quasar had a redshift of 4.73. The new one, PC1247+3406, has a redshift of 4.897 so that the light now reaching Earth from it was radiated when the Universe was only 7 per cent of its current age. Although it is not the brightest quasar yet seen, its image is well above the sensitivity of the astronomers' detector. This and the fact that the quasar was discovered after only 3 square degrees of the sky had been surveyed, suggests that we may hope soon to learn of even more distant objects.

1. Gherardi, E. & Stoker, M. *Nature* **346**, 228 (1990).
2. Weidner, K. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 7001–7005 (1991).
3. Stoker, M. *et al.* *Nature* **327**, 239–242 (1987).
4. Rosen, E. *et al.* *Cancer Invest.* **8**, 647–650 (1990).
5. Dowrick, P. G. & Warn, R. M. *Cytokine* **3**, 299–310 (1991).
6. Gherardi, E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 5844–5848 (1989).
7. Gohda, E. *et al.* *J. clin. Invest.* **81**, 414–419 (1988).
8. Nakamura, T. *et al.* *Fed. Eur. biochem. Soc. Lett.* **224**, 311–316 (1987).

9. Stoker, M. & Gherardi, E. *Biochim. biophys. Acta* **1072**, 81–102 (1991).
10. Nakamura, T. *et al.* *Nature* **342**, 440–443 (1989).
11. Naldini, L. *et al.* *Oncogene* **6**, 501–504 (1991).
12. Bottaro, D. P. *et al.* *Science* **251**, 802–804 (1991).
13. Cooper, C. S. *et al.* *Nature* **311**, 29–33 (1984).
14. Rosen, E. & Goldberg, I. *In vitro Cell dev. Biol.* **25**, 1079–1082 (1989).
15. Weidner, K. M. *et al.* *J. Cell. Biol.* **111**, 2097–2108 (1990).
16. Adams, J. C. *et al.* *J. Cell Sci.* **98**, 385–394 (1991).

CHEMICAL PHYSICS

Gas-phase liquids

Tony Stace

AT a qualitative level, the distinction between the spectroscopy of a gas-phase molecule and that of a molecular liquid can be summarized quite simply: isolated molecules yield very high-resolution spectra, whereas the spectra of liquids are poorly resolved because of the line-broadening processes which result from frequent collisions. But experiments by Sands *et al.*¹ now suggest that this distinction may not be as obvious as was previously believed. Results from experiments on the structure and dynamics of small, gas-phase clusters containing a

simple molecule and helium show that even when very cold (below 1 K), the inert-gas atoms continue to exhibit the type of large-amplitude motion normally associated with a liquid rather than a rigid complex.

Clusters can be formed by the adiabatic expansion of a high-pressure gas (20 atmospheres) through a small orifice (150 µm) into a vacuum (10⁻⁶ atmospheres). The rapid cooling (10⁸ K s⁻¹) that accompanies such an expansion process has two beneficial effects for spectroscopy: single molecules achieve very low rotational and vibrational temperatures, which results in a considerable simplification of spectra through a reduction in the range of possible spectral transitions; and condensation can take place which leads to the formation of complexes. Thus, by a suitable choice of gas/vapour mixtures, species such as He_nI₂, Ar_nSF₆, (H₂O)_nC₆H₆ (with *n* a small number) can be generated. The experiments of Sands *et al.*¹ rely on a combination of these factors in an attempt to measure the high-resolution spectra of He₂Cl₂.

In the spectroscopic analysis of isolated molecules like Cl₂ or C₆H₆, a first approximation is to assume that the molecules vibrate harmonically and that they remain more or less rigid as they rotate (the rigid rotor approx-

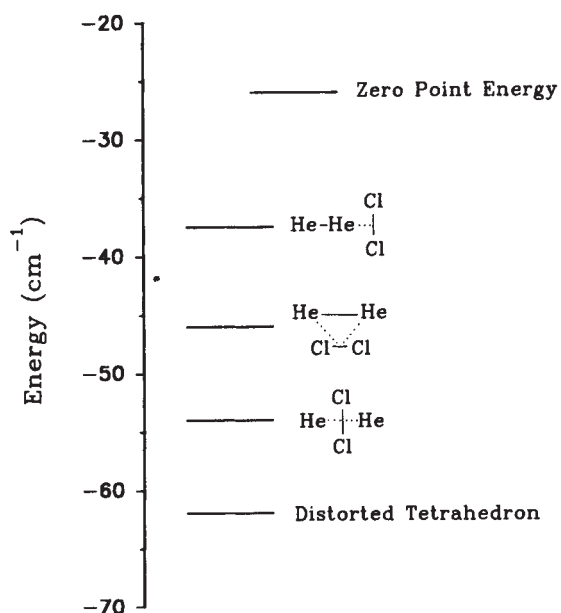


FIG. 1 Potential energy diagram for the possible configurations of the complex He₂Cl₂. Note that each configuration lies below the minimum energy (the zero-point energy) attainable by the complex.

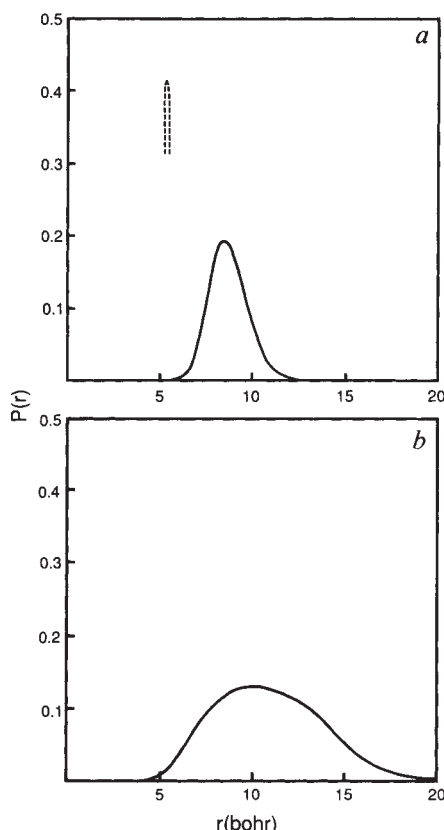


FIG. 2 The interatomic separation distribution for He-Xe (a) and He-He (b) in the ground state of He₂Xe. The broken line shows the distribution to be expected from a rigid rotor.

ximation). Thus, at temperatures close to absolute zero the relative positions of the atoms in a molecule or complex should be more or less fixed, with only the unavoidable zero-point vibrational motion contributing to any movement. From an analysis of spectra taken under these conditions, an equilibrium structure can be assigned. Earlier studies of Ne₂Cl₂ and Ar₂Cl₂ by Janda and colleagues^{2,3} showed that features in the spectra due to rotational motion could be accounted for by assuming that the clusters behave as rigid rotors. The difference with He₂Cl₂ is that, despite being below a temperature of 1 K, it does not follow the same pattern¹. Instead, the spectra show that, relative to the chlorine molecule, the helium atoms move considerably so that it is not possible to assign the complex an equilibrium structure.

The complex has four possible equilibrium configurations (Fig. 1) which can be calculated from a simple consideration of the interaction between the two helium atoms and between the molecule and each helium atom. The energy of each structure lies below the zero-point vibrational energy of the complex, so that even in their lowest energy state, the helium atoms have sufficient mobility to sample each structure on a time-

scale of several vibrational periods (10⁻¹² s).

Several research groups^{4,5} have attempted to calculate the dynamic properties of these and related systems (such as He₂Xe); a general conclusion appears to be that the complexes are best described as extremely floppy and liquid-like, and as having structural distributions rather than equilibrium configurations (Fig. 2). The broad spread of Xe-He and He-He interatomic distances shown in Fig. 2 for He₂Xe is characteristic of large amplitude motion over a very shallow potential-energy surface. A similar result would be expected for the He₂Cl₂ complex as it interconverts between the structures shown in Fig. 1. Taking the concept of a structural distribution one step further, one might liken Fig. 2 to one component of the pair or radial distribution function of a liquid; thus, as far as the spectroscopy of Cl₂He₂ is concerned, the core chlorine molecule is being perturbed by a continual sequence of (elastic) quasi-collisions with a liquid-like helium component.

From an experimental viewpoint, adiabatic expansion provides some (loose) control over the final temperature that can be achieved. Therefore, as mixed clusters of this type increase in size, it should be possible to identify either of two limiting physical states from their spectroscopy. With the heavier inert-gas atoms, such as argon and krypton, for which the interatomic interaction is comparatively strong, it seems that trapped molecules have spectra that are similar to those of matrix-isolated species⁶. Thus we would conclude that the clusters are solid-like. But by varying the expansion gas pressure, it has been suggested⁷, a phase transition can be promoted, whereby the clusters move from a liquid-like state to a solid-like one. Clearly, the phase diagram for bulk helium will permit only the liquid-like phase in even the largest of clusters. Interest then rests with the possibility that superfluid clusters may be generated⁸. □

Tony Stace is in the School of Chemistry and Molecular Sciences, University of Sussex, Brighton BN1 9QJ, UK.

1. Sands, W. D., Bieler, C. R. & Janda, K. C. *J. chem. Phys.* **95**, 729-734 (1991).
2. Hair, S. R., Cline, J. I., Bieler, C. R. & Janda, K. C. *J. chem. Phys.* **90**, 2935-2943 (1989).
3. Bieler, C. R., Evard, D. D. & Janda, K. C. *J. phys. Chem.* **94**, 7452-7457 (1990).
4. Horn, T. R., Gerber, R. B. & Ratner, M. A. *J. chem. Phys.* **91**, 1812-1823 (1989).
5. Garcia-Vela, A., Villarreal, P. & Delgado-Barrio, G. *J. chem. Phys.* **92**, 496 (1990).
6. Goyal, S., Robinson, G. N., Schutt, D. L. & Scoles, G. *J. phys. Chem.* **95**, 4186-4189 (1991).
7. Hahn, M. V. & Whetten, R. *Phys. Rev. Lett.* **61**, 1190-1193 (1988).
8. Toennies, J. J. in *Proc. CVII Summer School E. Fermi Int. School Phys.* (ed. Scoles, G.) 597-617 (North Holland, Amsterdam, 1990).

New vitality

LAST week, Daedalus proposed a 'life-cycle' for DNA. It survives the death of the organism itself (archaeological DNA millions of years old is known), and leaks back into the environment as stable drifting 'spores' of a few genes each. If such a spore happens to be present at an act of sexual fertilization, it can be incorporated into the new conception, and can start a new genetic career.

Over evolutionary time, quite a lot of spore-DNA must have managed to climb aboard the genome of each species in this way. Most of it will come from entirely different species. In its new home it will be so much meaningless parasitic junk — the 'junk DNA' so mysteriously present in most creatures. But a spore from the same species will change the creature's heredity. This is mutation.

Any dead tissue, says Daedalus, can release spores of DNA: even fallen hair, skin fragments, nail clippings and so on, discarded during life. A spore from a living 'donor' might easily find its way into some unrelated act of human love-making. The resulting child would then not be totally the offspring of its parents. A distant third party will have made a modest contribution.

The obvious application is to genetic engineering. If spores can be identified, and their structures understood, artificial ones could be constructed from specific regions of human DNA. The full mapping of the human genome will make it easier to choose regions which code for one precisely known characteristic. Introduced into human matings, they could produce small desired mutations to order. There must be an immune response to overcome, but this cannot be an insuperable barrier.

This minor tweaking of heredity raises far fewer problems than major interventions such as AID. Many potential parents might not be too jealous of a small controlled contribution from outside. Potential donors, too, would welcome the new technology. Even today, most childless people have twinges of regret; yet if the world's population is ever to decrease to a sensible size, many more of us will have to be childless. But with spore-DNA genetic engineering, anyone will be able to contribute a small but valued part of himself to other people's children, by donating snippets of hair or tissue for spore formation. Even after his death, his tissues could influence new generations. Physicists might happily incorporate a dash of the mathematical genius of Isaac Newton into their offspring; biologists could welcome a touch of darwinian imagination, while Lenin's tomb may prove a rich source of relic DNA for parentally hopeful politicians.

DAVID JONES

National greenhouse accounting

SIR — Various methods of estimating national contributions to global warming of emissions across different greenhouse gases have been proposed and discussed in anticipation of an international agreement aimed at slowing climate change. An understandable, comprehensive, technically based accounting method delineating national contributions to global warming is fundamental to a climate agreement. The method should reflect the relative warming ability of different greenhouse gases and assign responsibility for past, present and future contributions from different countries.

Two approaches to such an index have been proposed. The first, based on the work of Lashof and Ahuja¹ and adopted by the Intergovernmental Panel on Climate Change (IPCC)², creates a common potential warming unit, referred to as global warming potential (GWP), among the various greenhouse gases by integrating the unit radiative forcing over various future time periods for each gas relative to CO₂. The second, proposed by Hammond *et al.*³, avoids estimating the uncertain future radiative forcing of the GWP by examining an annual contribution to the atmospheric burden of greenhouse gases weighted by their instantaneous forcing relative to CO₂. This index has been called the global forcing contribution (GFC)⁴.

These approaches, although useful first steps, contain weaknesses which, when placed within the rubric of policy-making, may undermine efforts at agreement. I would like to propose a third method, based on an integrated forcing contribution, borrowing from the IPCC approach but emphasizing comprehensive national accounting in the distribution of global warming responsibility.

The national GFC of a particular greenhouse gas can be expressed as

$$\text{GFC}_{ij} = e_{ij} a_i / a_{\text{CO}_2} A_f \quad (1)$$

where a_i and a_{CO_2} are the instantaneous radiative forcing of a unit increase in the concentration of gas i and CO₂, respectively, and A_f is the airborne fraction of gas i . The airborne fraction can be simply expressed as

$$A_f = \Delta c_i / E_i \quad (2)$$

where Δc_i is the observed change in the concentration of gas i in the atmosphere in a given year, and E_i is the year's total global emissions of gas i .

The use of the airborne fraction, although attractive in relying on observable phenomena, has problems⁵⁻⁷. Gas removed from the atmosphere in a given year is mostly tied to emissions in pre-

vious years, not the year in which the airborne fraction is calculated. This means that if a nation's emissions in a given year were smaller than emissions in previous years, the airborne fraction would under-penalize that country. Furthermore, if the relative share of global emissions across all nations changed, the airborne fraction would produce unfair assignment of responsibility because the total atmospheric burden has been contributed to disproportionately in the past by all emitting nations. Finally, by using the observable change in the concentration of a greenhouse gas, non-anthropogenic emissions, shared out according to a nation's share of anthropogenic emissions, are included in the accounting and assignment of responsibility. Thus, the airborne fraction reflects past and present emissions but may assign responsibility for those emissions incorrectly.

Use of the GWP as an accounting method was suggested by the IPCC and can be expressed as

$$e_{ij} \text{GWP}_i = e_{ij} \frac{\int_0^\infty c_i(t) a_i dt}{\int_0^\infty c_{\text{CO}_2}(t) a_{\text{CO}_2} dt} \quad (3)$$

where e_{ij} is the emissions of gas i by country j ; $a_i(t)$ is the instantaneous radiative forcing due to a unit increase in the concentration of gas i ; N_f is the integration time horizon and $c_i(t)$ is the fraction of gas i remaining at time t which can be expressed as $c_i(t) = e^{-t/\tau_i}$, assuming the rate of removal is proportional to concentration. The parameter τ_i is the average lifetime of gas i , considered to be the time required to achieve a 1/e decrease in concentration.

Calculation of the GWP has been performed by the IPCC for 20-, 100- and 500-year integration time horizons. The GWP contains uncertainties, including quantifying the instantaneous radiative forcing, the lifetimes of various greenhouse gases and the difficulty associated with choosing an appropriate integration time horizon. As well as these, which have already been discussed^{1,8}, there is one crucial problem: The present and future radiative forcing from past emissions of greenhouse gases is left unaccounted for. Ignoring these past emissions will change the relative assignment of responsibility for greenhouse forcing, favouring large past emitters. The GWP itself is a useful quantity but should be incorporated into an expression that comprehensively accounts for past, present and future contributions to the burden of greenhouse gases in the atmosphere.

An alternative method, integrated forcing contribution (IFC), emphasizing

comprehensive assignment of responsibility, might be formulated as

$$\text{IFC}_i = C_{ij}(t) \text{GWP}_i \quad (4)$$

where $C_{ij}(t)$ is the concentration of gas i at time t resulting from emissions from country j . Rather than the use of CO₂ as a reference gas in the GWP, it would be best to use a gas with a more certain lifetime, such as CFC-11 or CFC-12.

$C_{ij}(t)$ can be expressed as

$$C_{ij}(t) = e^{-t/\tau_i} \int_0^t e^{t'/\tau_i} E_{ij}(t') dt' \quad (5)$$

which is a solution to the conservation of gas i in the atmosphere, where $E_{ij}(t)$ is the anthropogenic emissions from country j and boundary condition $C_{ij}(0) = 0$ applies. $E_{ij}(t)$ can be determined from historical data of national emissions or from records of energy use and land alteration, if direct emissions are not available. The data can be fitted to an emissions expression of the form

$$E_{ij}(t) = E_0 e^{r(t)t} \quad (6)$$

where E_0 is the emissions at time $t = 0$ and $r(t)$ governs the rate of change of emissions. This rate value can be expressed as a constant or include time dependence depending upon the nature of individual nation's emission patterns. Given an adequate statistical fit, incomplete data sets and data gaps can be overcome.

By quantifying past contributions to present greenhouse gas concentrations, and considering the present and future radiative forcing of those contributions, the IFC emphasizes a comprehensive accounting of national contributions to greenhouse warming. It avoids some of the weaknesses of the IPCC and GFC methods while including the necessary elements. Before considering the difficult but inevitable issues in creating an international climate convention, such as environmental equity, development rights and national control, an accounting method must be agreed upon. The IFC could provide additional framing of the relevant issues involved in creating such a method for the accounting and management of potential greenhouse warming.

KEVIN ROBERT GURNEY

*Tellus Institute,
89 Broad Street,
Boston, Massachusetts 02110, USA*

1. Lashof, D. A. & Ahuja, D. R. *Nature* **344**, 529 (1990).
2. Intergovernmental Panel on Climate Change *Scientific Assessment of Climate Change* (Cambridge University Press, New York, 1990).
3. Hammond, A. L., Rodenburg, E. & Moomaw, W.R. *Nature* **347**, (1991).
4. Handel, M. D. *Nature* **349**, 468 (1991).
5. Enting, I. G. & Rohde, H. *Nature* **349**, 468 (1991).
6. Victor, D. G., *Environment* **33**, 2 (1991).
7. Smith, K. R., *Environment* **33**, 4 (1991).
8. Victor, D. G., *Nature* **347**, 431 (1990).

Mimicry and viceroy butterflies

SIR — Ritland and Brower¹ reassess defensive mimicry among monarch, queen and viceroy butterflies in Florida (see Scientific Correspondence 351, 611–612; 20 June 1991). These three butterfly species have complex population fluctuations in both space and time in north Florida. The monarch, a long-distance migrant, is present in spring, breeding from April to June, and again on passage to overwintering locations during October and November². The experimental monarchs were collected in autumn, during their migration, at a time when their chemical defences are almost at their lowest effectiveness. By contrast, first-generation monarchs in spring have very high chemical defences and would produce predator responses much more like those described by J. V. Z. Brower³, or by Platt *et al.*⁴, rather than those reported by Ritland and L. Brower¹. This means that the spatial and temporal dynamics of monarch, viceroy and queen populations need to be assessed in relation to the dynamics and behaviour of predator populations.

How can we reach conclusions about the ecological operation of mimicry without understanding the diversity of potential natural enemies that might select for mimetic defences⁵? We need to assess which are the relevant natural enemies and then demonstrate how populations of models, mimics and natural enemies interact behaviourally, temporally and spatially and how their numbers change according to relative frequencies of interaction.

Moreover, Ritland and Brower studied butterflies at a well-known viceroy hybrid zone in Florida, which makes their conclusions even more difficult to accept. Even without these added changes in frequencies of different mimetic phenotypes, mimicry is complicated. The demonstration of unacceptability to one predator species is only one very small part of unravelling the dynamics of interacting populations of prey and predator complexes towards an understanding of mimicry.

STEPHEN MALCOLM

Department of Biology,
Imperial College, Silwood Park,
Ascot, Berkshire SL5 7PY, UK

1. Ritland, D. B. & Brower, L. P. *Nature* **350**, 497–498 (1991).
2. Malcolm, S. B. & Brower, L. P. *Experientia* **45**, 284–295 (1989).
3. Brower, J. V. Z. *Evolution* **12**, 32–47; 273–285 (1958).
4. Platt, A. P., Coppinger, R. P. & Brower, L. P. *Evolution* **25**, 692–701 (1971).
5. Malcolm, S. B. *Trends ecol. Evol.* **5**, 57–62 (1990).

SIR — Guilford¹ questions whether our experiment² really demonstrates that viceroys are unpalatable, suggesting instead that they could have been “gustatory mimics” of monarchs that the birds

previously tasted during the experiment. This interpretation is excluded because the six birds that received viceroys before monarchs or queens (birds A–F) showed aversions similar to the 10 birds that received viceroys after monarchs or queens (birds G–P in our Table 1)^{2,3}. Moreover, birds that found viceroys unpalatable did not necessarily treat monarchs or queens as unpalatable, and vice versa; no correlation existed either between the percentages of monarchs and viceroys eaten by individual birds, or between their rejection behaviours towards the two species³. These facts similarly exclude Guilford's interesting suggestion of pyrazine-based olfactory mimicry.

We agree with Malcolm's comment (above) on the spatiotemporal variation in monarch chemical defence, but this does not vitiate our finding that Florida viceroys are themselves unpalatable to an avian predator, with the attendant implications for mimicry theory. In addition, because we focused on palatability by feeding only abdomens to the birds, Malcolm's criticism that our viceroys were collected in a transition zone of wing colour phenotype seems irrelevant; the criticism also does not gain support from other data^{3–5}, indicating that viceroys outside the transition zone are also unpalatable.

DAVID B. RITLAND
LINCOLN P. BROWER

Department of Zoology,
University of Florida,
Gainesville, Florida 32611, USA

1. Guilford, T. *Nature* **351**, 611 (1991).
2. Ritland, D. B. & Brower, L. P. *Nature* **350**, 497–498 (1991).
3. Ritland, D. B. thesis (University of Florida, Gainesville, 1991).
4. Brower, J. V. Z. *Evolution* **12**, 32–47; 273–285 (1958).
5. Platt, A. P., Coppinger, R. P. & Brower, L. P. *Evolution* **25**, 692–701 (1971).

SIR — In the light of the seminal research done on the viceroy butterfly by D. B. Ritland and L. P. Brower (*Nature* **350**, 497–498; 1991) we have set up a group whose aim is to restore the proper place of this species in the scientific community and popular press. Obviously the viceroy should no longer be relegated to a vice-regal position with respect to the monarch. A committee will search for an appropriate common name, if one can use the term ‘common’ when referring to things regal. A number of suggestions have been made — sovereign, imperator, suzarin, emperor. Those with computer backgrounds thought that Viceroy Plus (or Viceroy +) would be appropriate, but others, concerned about the political future of Canada, thought that perhaps using the French word for monarch (*monarque*) would be better. There was then concern

about what each butterfly would be called in French translations, which are standard in Canada.

Our committee estimates that it will cost about Canadian \$10,000 to send a representative to the Nomenclature Committee of the World Lepidopteran Congress. To this end the local UP THE VICEROY committee will attempt to raise most of this through the sale of buttons, posters and T-shirts. The hottest selling money-maker is the UP THE VICEROY T-shirt, with two viceroys in the *Fighting Up* position. Any readers sharing our concern should send contributions to our treasurer, Dan Sarki, at our office in Thunder Bay.

WAYNE MACCALLUM
BRUCE THACKER
GORD JOHNSON

Lake Superior Fisheries Unit,
PO Box 5000,
Thunder Bay, Ontario,
P7C 5G6, Canada

Oil spill clean-up

SIR — On 24 March 1989, approximately 11 million gallons of crude oil spilled into the waters of Prince William Sound, Alaska, eventually contaminating nearly 1,000 miles of shoreline¹. Because of the apparent success of bioremediation on this oil spill, and the consideration of this technology for other oil spills², we have investigated the effect of bioremediation on ecology and human health.

To determine whether bioremediation would be effective, the US Environmental Protection Agency (EPA) conducted a large field demonstration project to evaluate the use of fertilizers for accelerating natural biodegradation of spilled oil by indigenous microorganisms¹. Both a slow-release, water-soluble fertilizer and an oleophilic formulation were used for the treatment of the beaches.

During biodegradation, the toxicity of the polluting material may increase when non-toxic constituents are converted to toxic species. Toxic compounds refractory to biological degradation may further compromise the effectiveness of the treatment technique. In such complex situations, it is impractical to use analytical chemistry to detect and quantify all toxins that have the potential to cause chronic effects. Consequently, biological test systems complement chemical-specific analyses. Because carcinogenic and mutagenic effects are of interest both to environmental and to health scientists, the EPA bioremediation efforts were monitored with the spiral *Salmonella*/mammalian microsome mutagenicity assay³. The *Salmonella* assay has been used to demonstrate the mutagenicity of crude oils, including

Prudhoe Bay crude oil and the correlation between mineral oil carcinogenicity and *Salmonella* mutagenicity⁴. The mutagenicity of the bioremediation products of crude oils, however, has not previously been examined.

In our study, the *Salmonella* mutagenicity assay showed that both the Prudhoe Bay crude oil and its weathered counterpart collected from oil-impacted water are weakly mutagenic, but that neither of the fertilizer formulations are mutagenic. When examined on the basis of mutagenic activity per amount of soil, the sand/gravel beach had higher mutagenic response than the cobblestone beach. The difference in the beach types seems to be related to the total surface area available resulting from differences in particle size. The mutagenicity of samples collected from the contaminated beaches declined over time as a result of both bioremediation and natural processes, and the mutagenic components were depleted faster than the overall content of organic material.

Because oil spills involve dynamic processes and differ in the amount and distribution of oil contamination, the microbial ecology, the chemistry of the oil, weather conditions, and so on, the results of this study cannot be extrapolated to attempts elsewhere to use bioremediation to deal with oil spills. This study does indicate, however, that both natural and enhanced bioremediation processes can degrade the mutagenic components of crude oils.

LARRY D. CLAXTON
VIRGINIA S. HOUK

Health Effects Research Laboratory,
US Environmental Protection Agency,
Research Triangle Park,
North Carolina 27711, USA

RON WILLIAMS

Environmental Health Research and
Testing,
Research Triangle Park,
North Carolina 27709, USA

FRAN KREMER

Center for Environmental Research,
US Environmental Protection Agency,
Cincinnati,
Ohio 45268, USA

1. Pritchard, P. H. & Costa, C. F. *Environmental Sci. Technol.* **25**, 372–379 (1991).
2. Aldhous, P. *Nature* **349**, 447 (1991).
3. Houk, V. S., Early, G. & Claxton, L. D. *Environ. molec. Mutagen.* **17**, 112–121 (1991).
4. Blackburn, C. R., Deitch, R. A., Schreiner, C. A. & Mackerer, C. R. *Cell Biol. Toxicol.* **2**, 63–84 (1986).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

Sea-level curve

SIR — We have obtained a carbon-14 date for charcoal in an ancient fire-pit, providing an independent check on the sea-level record reported by Fairbanks¹. This charcoal, now 27.4 m beneath the local groundwater table, gives a ¹⁴C age of $8,250 \pm 80$ yr (uncorrected for secular changes in atmospheric ¹⁴C). The fire pit is a cavity in a limestone block resting on the irregular floor of a cave (incongruously named "Car Wash" cenote) near Tulum, Quintana Roo, Mexico. We dated particles of charcoal ranging from clay size to 1 cm in maximum dimension: 500 ml of charcoal-water slurry containing 2.9 g (dry weight) charcoal yielded 860 mg charcoal after purification. The water table in the area is nearly flat; it is unlikely that the water level in Car Wash was ever more than a metre above contemporary sea level^{2,3}. The groundwater here is a lens of fresh water floating on a saline intrusion⁴ with a fresh/salt water interface at about 20 m — thus, the fire pit is in salt water. The charcoal age gives the oldest date at which the cave could have remained dry to a depth of 27.4 m, hence, the oldest date in the past 20,000 years when the sea level could have been as much as 28 m below its present mean value.

Fairbanks' glacio-eustatic sea-level curve¹ is based on ¹⁴C dating of coral (*Acropora palmata*) from the Atlantic Ocean. Our date, which gives a maximum sea-level estimate, agrees well with Fairbanks' minimum level estimated for that time. Both Fairbanks' data and ours are sensitive to secular changes in atmospheric ¹⁴C (not corrected for). Fairbanks also assumed that the 400-year offset between surface ocean radiocarbon and atmospheric radiocarbon did not change throughout the time range of interest. We assume that the charcoal in our sample derived from terrestrial plants of the same age as the fire pit. Fairbanks applied a 34 cm per kyr correction for uplift; we have not applied a tectonic correction.

This limited comparison indicates that standard terrestrial and marine ¹⁴C dates are closely comparable and that two separated areas (Barbados and the Yucatan peninsula) suffered little, or much the same, vertical tectonic movement in the past 8,500 years.

Car Wash cenote is 40 m south of the Coba-Tulum highway, 8 km west of Tulum. The fire pit, discovered in 1986 by M. Madden and P. Turner, is 130 m from the cave entrance, and, as illustrated in ref. 6, is well above floor level. Charcoal found on the cave floor is consistent with human occupation. Our age/depth assignment holds if the charcoal was deposited before the cave filled with water. The high proportion of char-

coal to mineral particles and the wide variation in particle size (settling time) of the charcoal make it unlikely that it was introduced by flowing water.

JAMES COKE

Postal 1,
Playa del Carmen,
Quintana Roo, Mexico 77710

E. C. PERRY

Department of Geology,
Northern Illinois University,
DeKalb, Illinois 60115, USA

AUSTIN LONG

Department of Geosciences,
University of Arizona,
Tucson, Arizona 85721, USA

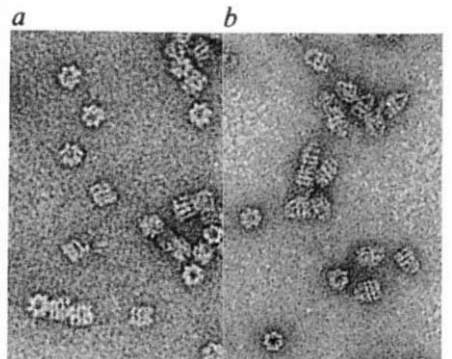
1. Fairbanks, R. G. *Nature* **342**, 637–643 (1989).
2. Back, W. & Hanshaw, B. J. *Hydrol.* **10**, 330–368 (1970).
3. Martin, L. E. thesis (Northern Illinois University, 1990).
4. Stoessell, R. K., Ward, W. C., Ford, B. H. & Schuffert, J. D. *Geol. Soc. Am. Bull.* **101**, 159–169 (1989).
5. Lighty, R. G., MacIntyre, I. G. & Stuckenrath, R. *Coral Reefs* **1**, 125–130 (1982).
6. Lockwood, C. C. *The Yucatan Peninsula* (LSU, Baton Rouge, 1989).

Binding of chaperonins

SIR — A model for the action of groE proteins in assisting protein folding was presented by Creighton in News and Views¹, based on the stoichiometry of one groEL 14-mer to one groES 7-mer². GroES is an essential cofactor for the groEL-mediated folding of at least some proteins³.

Creighton made a detailed proposal for nonsymmetric binding of groES to one face of the groEL 14-mer, and ended with the remark that a verdict on this and other models could soon be expected. Some of the pertinent data already exist and have been presented in public. We have been studying the groES–groEL complex, which we have crystallized and also observed by negative-stain electron microscopy, and which bears directly on the question of asymmetric binding.

The figure shows top and side views of *Escherichia coli* groEL (a) and the groEL–groES complex (b) prepared in the presence of ATP. The characteristic four-striped side view, in which the 7-



Negative stain of groEL (a) and groEL–groES complexes (b). Magnification, $\times 189,400$.

fold axis probably runs perpendicular to the stripes⁴, is markedly perturbed by the binding of groES. One of the previously flat surfaces is bowed out, forming a dome in the side view. This clearly shows the asymmetry of groES binding, and suggests a structural change in one disk of the groEL oligomer to produce an increase in volume greater than would be caused by the simple addition of the M_r 70,000 groES 7-mer to the M_r 840,000 groEL. This may be an indication of conformational changes in groEL suggested to occur as part of ATP-driven protein folding³.

HELEN SAIBIL
ZHENG DONG
STEVE WOOD

ARLENE AUF DER MAUER

Department of Crystallography,
Birkbeck College,
Malet St, London WC1E 7HX, UK

1. Creighton, T. E. *Nature* **352**, 17–18 (1991).
2. Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R. & Georgopoulos, C. *J. biol. Chem.* **261**, 12414–12419 (1986).
3. Martin, J. *et al.* *Nature* **352**, 36–42 (1991).
4. Hutchinson, G., Tichelaar, W., Hofhaus, G., Weiss, H. & Leonard, K. R. *EMBO J.* **8**, 1485–1490 (1989).

Stem cells

SIR — Jones *et al.* have demonstrated¹ that cells can be separated from mouse bone marrow which contribute 50–80% of the DNA in myeloerythroid, B and T cells 60 days after irradiation, but which possess no radioprotective and little CFU-S activity. They conclude “that there are two vital classes of engrafting cells: committed progenitors, which provide initial, unsustained engraftment, and pluripotent haematopoietic stem cells, which produce delayed, but durable, engraftment.” Iscove in his News and Views article² contrasted these results with our report³ that Thy-1^{lo} Lin[−]Sca-1⁺(Th^{lo}L[−]S⁺) cells representing only about 0.05% of all bone marrow cells, are 1–2,000-fold enriched in radioprotection, CFU-S (day 12), and long-term multilineage reconstituting (LTMR) activity. Iscove concluded “the extent of possible overlap of day-12 CFU-S and long-term reconstituting units appears to be very low indeed and seems incompatible with conclusions drawn by Spangrude *et al.*”.

Iscove's comments could be taken to mean that the Th^{lo}L[−]S⁺ cell type we described³ is not the most highly enriched fraction of haematopoietic stem cell yet obtained, or that the fraction does not even contain haematopoietic stem cells by the definition of LTMR. That is not so. We have shown that 1–20 Ly-5 marked Th^{lo}L[−]S⁺ cells mixed with 100 unmarked Th^{lo}L[−]S⁺ cells will allow the assay of a single clonogenic event *in vivo*⁴. In most cases, marked clones were multi-lineage (T, B, myeloid). About

30% of the clones gave sustained engraftment for more than 9 weeks. The rest stopped producing new progeny before the 9-week time point. These two outcomes could result from a primary heterogeneity in the stem cells, or a chance event such as location in a self-renewing microenvironment. We tested⁵ whether the subset could be divided into subsets using the dye Rh123. Rh123^{lo} cells contain both measurable day-12 CFU-S and were highly enriched for LTMR, whereas Rh123^{hi} cells contained fourfold greater CFU-S day-12 activity than the Rh123^{lo} cells, but were significantly less effective at giving rise to pre CFU-S and LTMR.

We believe that any differences between the results of Jones *et al.* and ours are likely to be obscured by the problems of studying a highly enriched, extremely rare population compared with studying populations in which the separation technology is not designed to isolate rare subsets, and which may either alter cellular potentials or contain regulatory cells.

Iscove concluded “for the cells of greatest practical interest — those with long-term reconstituting ability — there still seems to be no satisfactory substitute for long-term assays *in vivo*.” We agree and have done such assays^{3–6}. If we assume that Iscove means by the term “practical” that one principal objective of the isolation of haematopoietic stem cells is to restore lethally irradiated or myelo-ablated humans, we would say that long-term reconstitution alone is not sufficient: Jones *et al.* reported that none of 16 mice survived with 1×10^5 25 ml per minute cells (the fraction capable of long-term reconstitution), certainly an unacceptable outcome for bone marrow transplantation in man.

IRVING WEISSMAN
GERALD SPANGRUDE
SHELLY HEIMFELD
LAURIE SMITH
NOBUKO UCHIDA

Beckman Center for Molecular and
Genetic Medicine,
Stanford University Medical Center,
Stanford, California 94305, USA

ISCOVE REPLIES — Weissman and colleagues could mean that they consider the Thy-1^{lo}Lin[−]Sca-1⁺ fraction of mouse bone marrow to be purified to a similar degree in both long-term multilineage reconstituting stem cell (LTMRSC) activity as well as CFU-S. The issue is important: since their fraction is nearly pure in CFU-S, near purity of LTMRSC would also be implied — the end of the road, as it were. In the News and Views piece in question², I pointed out the apparent discrepancy between that interpretation and the findings of Jones *et al.*¹, showing virtually complete physical

separation of CFU-S and radioprotective activity away from the bulk of the LTMRSC in marrow. If correct, the numbers were strong enough to imply that the majority cell type in the Th^{lo}L[−]S⁺ fraction (CFU-S) had to be something other than LTMRSC. The discrepancy was more than trivial, and to explain it away as “likely to be hidden in the quantitation of studying a highly enriched, extremely rare population. . .” seems less than sufficient.

In support of the claim for extensive co-purification of LTMRSC in Th^{lo}L[−]S⁺ fraction, Weissman *et al.* cite their more recent study⁴. In every case of successful clonal reconstitution achieved with their fraction, they write, the response was multilineage, “in 20–35% of the cases the clone gave sustained. . . engraftment” and, in the latter case, “retransplantation. . . was the rule”. However, the statements lose in impact when the actual numbers are examined. Only 1/40–1/13 Th^{lo}L[−]S⁺ cells yielded detectable blood cell production up to 3–7 weeks, fewer continued much longer, and in only a solitary instance (mouse C7) of 280 mice transplanted with purified cells was reconstitution durable enough to survive transfer into secondary hosts.

If nothing else, Smith *et al.*⁴ and Spangrude and Johnson⁵ confirm that the Th^{lo}L[−]S⁺ fraction is considerably more heterogeneous than was originally appreciated. LTMRSC may be present, but the experiments to date establish their presence only as a very small minority. The presence of a few is not terribly informative. What does matter is how they *distribute* in fractionation experiments, and whether the presence of a few in the Th^{lo}L[−]S⁺ fraction could represent an experimental contaminant whose true phenotype could differ significantly from that of the bulk of the fraction. Reconstitution assays have been available for some time for measurement of LTMRSC specifically, quantitatively and independently of short-term radioprotection (see, for example, ref. 7). Appropriate application of such assays will clarify the current uncertainties and one can only look forward to publications in the refereed literature which do so.

NORMAN ISCOVE

Ontario Cancer Institute,
500 Sherbourne St,
Toronto, Canada M4X 1E9

1. Jones, R. L., Waner, J. E., Celano, P., Zicha, M. S., & Sharkis, S. J. *Nature* **347**, 188–189 (1990).
2. Iscove, N. *Nature* **347**, 126–127 (1990).
3. Spangrude, G. J., Heimfeld, S. & Weissman, I. L. *Science* **241**, 58–62 (1988).
4. Smith, L., Weissman, I. L. & Heimfeld, S. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2788–2792 (1991).
5. Spangrude, G. J. & Johnson, G. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7433–7437 (1990).
6. Ikuta, K. *et al.* *Cell* **62**, 863–874 (1990).
7. Harrison, D. E. *Blood* **55**, 77–81 (1980).

Java on the brink

Jared Diamond

Zoological Researches in Java, and the Neighbouring Islands. By Thomas Horsfield. Oxford University Press: 1990. No pagination. £175, \$350.

"YOUR industry, sir, in collecting them is praise-worthy in the extreme; and the talent you have shown in arranging them encourages a well-founded hope of much advantage to science being derived from your arrangement and observations on them." This compliment, written by Sir Joseph Banks around 1817, was addressed to an American physician named Thomas Horsfield, who spent 18 years collecting and describing animal and plant specimens ("them") on the Indonesian island of Java. Between 1821 and 1824, Horsfield published a lavishly illustrated account of 64 Javan bird and mammal species. That rare landmark in Asian zoology has now been reprinted with an introduction by John Bostin that reconstructs Horsfield's journeys — a task made necessary by Horsfield's having ordered the executor of his will to burn all his diaries and field notes.

As an American working in one of the battlegrounds of Far Eastern colonialism, Horsfield profited from all three European powers involved in the struggle. On arriving in Java when it was a Dutch colony, he was hired to survey Java's poisonous and medicinal plants. He was kept on by the British when they defeated the Dutch in 1811, and was rehired by his former employers when Java was returned to Holland in 1816; but he continued to send his collections to London's East India Company Museum, of which he became a curator on leaving Java. His career was advanced by the British colonial administrator Sir Stamford Raffles in many ways, such as by Raffles' seizure and shipment to London of a large zoological collection amassed by French naturalists. Fearing that the French might nevertheless publish first, Raffles prodded Horsfield to beat them to it. The resulting book is interesting as an account of a remarkable fauna, as a developmental stage in the nineteenth-century tradition of illustrated natural-history treatises, and as a document in conservation biology.

With regards to Java's fauna, the word "island" is often coupled with the phrase "faunally impoverished". Java, however, acquired a rich fauna owing to its equatorial location and high mountains (up to 12,000 feet), as well as its land connections to Asia that existed during the Pleistocene epoch, when sea-level was low. Over those now-flooded land bridges, Java man (*Homo erectus*)

walked to Java nearly a million years ago, along with representatives of many other tropical Asian species. Alfred Russel Wallace's observations of these species in Java and Bali, which were absent on nearby Lombok, stimulated him to recognize the world's most striking faunal transition (Wallace's line) and then to formulate the modern science of



Endangered species — *Falco caerulescens*, at around six and a half inches long, is thought to be Java's smallest Falcon.

biogeography. Among the Javan species that Horsfield selected for his book are many of these Asian specialties, including a rhinoceros, tapir, gibbon, leafbird and fairy bluebird.

Horsfield's book belongs to the tradition of lavishly illustrated, large-format, limited-edition natural-history volumes that reached their apogee soon afterwards in John James Audubon's *The Birds of America* and John Gould's many sets of bird books. Today these books are not only rare but expensive, selling at auction for up to millions of dollars. Horsfield's was the second British book to contain hand-coloured lithographs of birds, a practice that became standard with his successors (including Audubon and Gould), who greatly improved on the quality of Horsfield's production.

While a few of Horsfield's bird and mammal illustrations are lively and

beautiful, many have unnatural proportions or postures. The text consists mainly of descriptions of anatomy and external appearance, with usually only brief accounts of the behaviour and distribution of the fauna. Lacking is Wallace's interest in broader questions, such as how flightless mammals came to be present on an island.

Instead, Horsfield's lasting value to zoologists rests on his early exploration of a vanishing biota. His collections were the first large, systematic, natural-history collections from Indonesia to reach Europe. He was able to list nearly half of the bird species subsequently recorded from Java. Today, Java is one of the most overpopulated areas of the

world, with over 100 million people jammed into an area equivalent to Greece. One-third of Java's resident bird species are endangered, and a dozen of them may already be extinct. The reprinting of Horsfield's book is therefore useful not only to bibliophiles interested in the evolution of illustrated natural-history books, but also to conservation biologists for whom Horsfield offers a poignant snapshot of species on the brink of extinction. The reprint is indeed expensive, but the original is simply unavailable except at auction, and then only to millionaires. I hope that presses will now be inspired to reprint the many other unaffordable illustrated natural-history books of comparable scientific interest. □

Jared Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024, USA.

Struggling for air

Robert Temple

History of Science and Technology in Ancient India, Vol. II: Formation of the Theoretical Fundamentals of Natural Science. By Debiprasad Chattopadhyaya. Firma KLM Private Ltd, 257B, Bepin Behari, Ganguly Street, Calcutta 700012, India: 1991. Pp. 593. Rupees 625.

THIS is more than a book on the history of science. It grapples directly with the issue of whether India is to have any future or not. Chattopadhyaya is a brave man, and he has tackled the fundamental problem head on: he shows the history of Hindu obscurantism that has suppressed the rise of science in India through the ages — the implications for the present are clear.

There were many brilliant early scientists in India, and it is in this book that we learn of them essentially for the first time. The author's historical spadework is breathtaking. He reconstructs the true story through the fog of the intervening religious fanaticism, and undoes the tangled knots of mangled texts brought about by the centuries of distortion and suppression. Westerners will be amazed to learn that human corpses were routinely dissected in India centuries before Vesalius (1514–64), as part of what Chattopadhyaya rightly calls ancient Indian "rationalist medicine". The relevant ancient text specifically calls for "direct personal observation". The original form of Indian Ayurvedic medicine was thoroughly and spectacularly rational. But its sad distortion, overlay and indeed perversion by religious fanatics has necessitated the use of Chattopadhyaya's own surgical skills in extracting the original glories of ancient Indian medicine from the mounds of rubbish in which it became embedded.

This is one of the saddest books ever written about the history of science. For never has a culture so satisfactorily stifled scientific progress as Hindu culture. The smug self-satisfaction of the devout — and they nearly won a recent election — has put a wet blanket over generation after generation of brilliant men of science. The Indian genius is there; but so is its nemesis. This book deserves to be read as a case history of how rationalism can be defeated repeatedly over the course of three millennia. There is no parallel in the annals of human thought.

Chattopadhyaya convincingly shows at great length that the scientist Uddālaka aśruni antedated Thales of Greece and

Wounded civilization

AYURVEDIC medicine is in many ways still the medicine of millions. Yet it seems that all it had to offer the authors of the Hippocratic books was pepper for the eyes and a mouth-wash for bad breath. As a system of medicine literally meaning 'knowledge of life', Ayurveda was incorporated into Hindu tradition as part of the fourth Veda, one of the ancient sacred books in Sanskrit. The system was based on two treatises composed in verse and prose by Charaka and Sushruta. Their dates, as with those of ancient Indian history in general, are woefully uncertain. Sushruta, for example, could have lived anywhere between 1000 BC and AD 1000. In *The Healing Hand: Man and Wound in the Ancient World*, Guido Majno returns to such original sources to unravel the history of the physician's art in the ancient civilizations of India, Greece, Rome, Egypt and China. With considerable verve, he reconstructs



case histories to see whether the patients were made better or worse by the treatments, often putting ancient remedies to the test in his own laboratory. In one section he assesses the advice given by Sushruta to physicians who run into complications with pierced or stretched ears, a practice widespread in India owing to the protection it supposedly gave against "evil influences of malignant stars and spirits". This fifth- or sixth-century cave painting shows the stretched earlobes of a celestial nymph. In cases of "extreme pain and swelling", Sushruta's advice was to remove immediately the lint that stuffed the hole, anoint the infected part with a paste of honey and clarified butter enriched with barley and four herbs, wait for it to heal, then start all over again. Majno's panoramic book pinpoints beautifully those rare occasions in history when one great ancient civilization reaches out to learn from another. The book is to be issued in paperback by Harvard University Press on 26 September. Price is \$19.95, £15.95.

PT

that Indian science is older than Greek science. In the seventh or eighth century BC, Uddālaka held to a position of "essentially a rational knowledge based on observation and even experiment". He believed that "everything is made of indivisible partless entities". He conceived "the basic stuff of the world as consisting of finest essence, of invisible minute particles". But more important, "instead of just proclaiming some theory or seeking any scriptural declaration as its main prop, he wants to demonstrate it on the strength of actual observation at every step".

In no other culture did atomism have a longer history. As Chattopadhyaya points out, "in the Indian tradition atomism was much more than a somewhat passing episode as it was in ancient Greece." He says: "The defence of atomism therefore specially in the nature-philosophy of the Nyāya-Vaiśeṣikas stretched over many centuries seems to be a unique feature of the scientific tradition of ancient India."

This is in stark contrast to China, where an atomic hypothesis never took root at all. As recently as the seventeenth century in India, the philosopher Gadadhāra on his deathbed uttered as his last words "Pīlu, pīlu, pīlu", which means "Atoms, atoms, atoms". Thus, atomism was a traditional Indian hypothesis for more than 2,200 years.

This superb work of scholarship defies all summary. The author has rendered a service to all historians of science by opening up an unknown world — unknown even to his fellow countrymen. Let us hope that the lessons will not be lost on them.

If India continues to allow religion to have the upper hand over science, then the tales told by Chattopadhyaya will have sequels, and India will relapse into the Stone Age. That, frankly, is his message. □

Robert Temple is at David Higham Associates Ltd, 5–8 Lower John Street, Golden Square, London W1R 4HA, UK.

The fish that time forgot

Michael A. Taylor

Living Fossil: The Story of the Coelacanth. By Keith Stewart Thomson. Norton/Radius: 1991. Pp. 252. \$19.95, £15.99.

THERE are indeed special reasons for the worldwide fascination with the 'living fossil' fish *Latimeria chalumnae*, better known as the coelacanth. All the same, I am sure Keith Thomson suffers from the coelacanth strain of the syndrome that afflicts my research on plesiosaurs. Any mention of these extinct marine reptiles too often triggers an enquiry about the Loch Ness monster. With age and wisdom, I no longer damn the Scottish Tourist Board, but instead discuss the plausibility of *Nessiteras*. This is a surprisingly severe test of anyone's knowledge of plesiosaurs, and gives the layperson a demonstration of science in messy action. And, not least, it wins half the battle by capturing the audience's attention from the start.

Thomson has the same advantage, with the added benefit that biologists have managed to catch living coelacanths, although they are still at a loss to explain why the fish should exist at all. This enjoyable look at coelacanths will suit almost everyone from interested laics, through students wanting background reading, to any zoologist or geologist who is not a sarcopterygian specialist.

Thomson's measured and understated, but still personal, style has been honed in the essay column of *American Scientist*, and is a refreshing contrast to the personal aggrandizement and even backbiting of some other writers on evolution and fossils. He does not expand unduly on his own research. When he tells how he obtained the first frozen coelacanth (much more informative than the pickled version) by simply asking the Comoran government for it, he can only wonder at his naivety.

The first-recorded catch of the coelacanth was made off the coast of southern Africa in 1938. The discovery was announced in a short notice to *Nature*, albeit not without being scooped by the *London Illustrated News*. Disputes continued with the capture of the second specimen off the remote Comoro Islands, then the territory of the French government, who claimed a legal right to the fish. After the capture of the third specimen in 1953, the French declared a complete ban on expeditions in the area by foreign scientists, which was not lifted until the end of French colonial rule

some 15 years later. There was even a ban on dissecting any specimens sent to foreign museums.

The key discoveries were made by people who were unsupported by conventional science funding, but who nevertheless knew an opportunity when they saw one: Marjorie Courtenay-Latimer, who, as curator of a tiny museum in Cape Province, South Africa, discovered the fish, recognizing it as something odd; J. L. B. Smith, the chemistry lecturer and spare-time ichthyologist who identified it as a coelacanth and wrote the paper to *Nature*; and Hans Fricke who raised private cash

there. A few more illustrations and labels in the book might have helped explain matters such as notochords and humeral homologies, but most readers will get enough succinct grounding in topics such as the nature of living fossils, continental drift, the origin of tetrapods, and taxonomy, to appreciate *Latimeria's* significance. I certainly came away fully appreciating that the many gaps in our knowledge of *Latimeria's* lifestyle are nothing compared with the puzzle of how it apparently came to live only in the impoverished habitat of a few underwater lava slopes. Moreover, these slopes have existed for merely a fraction of the roughly 70 million years that separate the fish from the last-known fossil coelacanths.

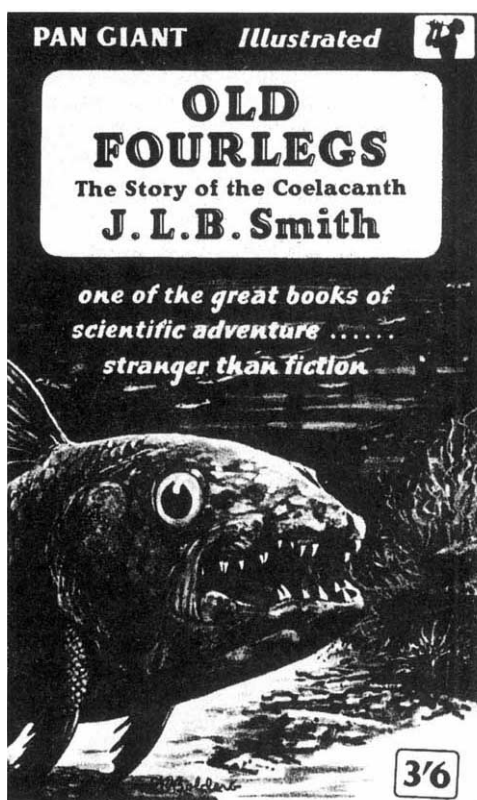
Thomson reveals the humanity of scientists and the fallibility of memory. Nevertheless, his hero remains *Latimeria*, so read the book, then watch Fricke's extraordinary documentaries of coelacanths in the wild. In these, *Latimeria*, like Lewis Carroll's Father William, comes across as an old if sometimes inscrutable member of the family. The two share many traits: *Latimeria* will incessantly stand on its head, keeps all its limbs very supple and swears by large amounts of oil, at least for buoyancy control. But instead of balancing an eel on its nose, it has the amazing, apparently electroreceptive rostral organ. Despite the coelacanth's nickname of 'old fourlegs', its remarkably mobile paired fins are a drift-hunter's means of staying on station in a slow current, rather than the bottom-crawling fins of tetrapod ancestors. So 'father' seems purely honorific. As Thomson stresses, *Latimeria* is "not trying to be an amphibian or an amphibian trying to be a fish. It is a perfectly good fish. . .".

Unfortunately, *Latimeria* may never recover from our obsessive interest in our close relative (but, as Thomson points out, not *that* close).

As J. L. B. Smith long ago feared, its slow reproductive rate and our utter ignorance of its true population size mean that every capture of a female — whether for research, commercial exhibition or tourist souvenir — could be driving the species to extinction. Captive breeding might well destroy, and will at least alter, the very species it aims to save. Leave the fish alone, Thomson says, for if *Latimeria* becomes extinct, there really will be no one else to blame but us scientists.



Michael Taylor is in the Earth Sciences Section, Leicestershire Museums, Arts and Records Service, 96 New Walk, Leicester LE1 6TD, UK.



The cover of J. L. B. Smith's popular book (1956). The picture is one of many in *The Biology of Latimeria chalumnae and evolution of coelacanths* (eds J. A. Musick, M. N. Bruton and E. K. Balon), the largest and most comprehensive volume so far published on the fish. Published by Kluwer, price is Dfl. 350, \$199, £177.50.

for his Submarine photographic observations. But they lived in relative wealth compared with the fishermen such as Ahmed Hussein Bourou who did the hard work from tiny pirogues with handlines hundreds of metres long. Today there is practically uncontrolled exploitation of coelacanth fishing by the now independent and desperately poor Comorans.

This is not a book of academic minutiae, whether about the history of science or fish evolution. Thomson seems to have decided that detail such as the new Greenland fossil tetrapods is irrelevant to the central theme of the coelacanth, how it lives and breeds, and why it is

Change in the scenery

Michael Church

Global Geomorphology. By Michael A. Summerfield. Longman: 1991. Pp. 537. £17.99.

GEOMORPHOLOGY ought to be the most interesting part of the Earth sciences. Landscape is accessible, and most people have at least a passing interest in how scenery came to be the way it is. Moreover, human societies draw nearly all of their resources from the surface of the land, and build nearly all of their structures there. But by the time the idea of plate tectonics burst upon us 30 years ago, geomorphologists had already begun to be preoccupied with the mechanics of erosion on local scales, and so the subject became a backwater.

Nevertheless, the founders of the modern subject — pre-eminently W. M. Davis and L. C. King — were certainly aware of the essential interplay of tectonics, erosion and sedimentation in creating landforms. Recently, geomorphologists have started to regain this perspective. *Global Geomorphology* makes the most determined attempt yet to exemplify it in a textbook. Accordingly, the book is divided into three principal sections concerned with endogene processes (read 'tectonics'), exogene processes (erosional development of the landscape), and endogene-exogene interactions. In the last of these, emphasis is given to rates of uplift and denudation, to tectonic control of drainage and sea level, and to long-term landscape development. The plan of the book is stimulating.

But the book is far less successful at surmounting some other important issues that have dogged the subject. Geomorphology is not highly quantified, mainly because at the scale of large landforms it is dominated by complex boundary conditions, making it difficult to give an adequate analytical representation of the phenomena involved. Furthermore, geomorphologists deal chiefly with one Earth and its history: sequences of events often appear to be unique, and the number of tolerably similar phenomena is frustratingly small. Summerfield gives an illuminating discussion of these issues in his opening chapter (although I doubt that inexperienced students will appreciate it properly), but sublimates them in the rest of the book by adopting a completely verbal and descriptive manner. The author claims that a mathematical treatment would simply not be appropriate for most students. But this compromise

demotes the book to being a description for nonspecialists, rather than the important summary of the field that it might otherwise have been.

A verbal argument can be exciting if it conveys ideas and arguments in a lucid and provocative way. But Summerfield is so preoccupied with definitions and classifications that basic principles are obscured or slighted. In the discussions of tectonic effects, he emphasizes the difficulties in attempting to reduce to simple models the features of complex, unique landscapes in which history and configuration matter a lot. But the reading is heavy-going because there is no unifying thread. Even the generous number of pictures and diagrams, borrowed from many sources, reinforce the sense of there being too many details and too little essential explanation. In the more traditional sections on exogene processes, the obverse problem arises. The discussion fails to illustrate that geomorphological processes are complex physical and chemical phenomena. The forces on slopes, fluid flow and sediment trans-

port — the fundamental phenomena in the formation of most landscapes — remain inadequately characterized.

The book is intended to be read by undergraduates who have a basic knowledge of the physical environment. For them its outstanding feature is that it collects all its references into short bibliographic essays at the close of each chapter. The choice of references is especially good. Paradoxically, though, the book may be of most value to more advanced students, for whom it presents some interesting themes and questions. It re-emphasizes the history of individual landscapes, focuses on the interaction of tectonics and erosional development of landscape, and questions whether the geomorphology of other planets should become a larger part of the traditional subject matter. The author has a worthwhile vision of the field, but unfortunately has not presented it well. □

Michael Church is in the Department of Geography, University of British Columbia, Vancouver, Canada V6T 1Z2.

No more, no less

B. C. Jarvis

Plant Physiology. By Lincoln Taiz and Eduardo Zeiger. Addison-Wesley: 1991. Pp. 559. £21.95.

THE problem facing authors of general texts such as this is not so much what to include, but rather what to leave out. The objectives of Taiz and Zeiger are admirable — to outline the principles of physiological investigations and to discuss a wide array of physiological processes in plants, while attempting "to convey the excitement and direction of contemporary research". To these ends they have elicited contributions from 21 other researchers, and have achieved a consistency of style and clarity.

The book begins with a 'primer' that describes plant and cell architecture and examines fundamental concepts about energy, enzymes and gene expression. Basic it may be, but it is essential background material. A section on the control of metabolism *per se*, however, would have added to its usefulness.

The remaining 19 chapters are grouped conventionally into three units: 'Transport and Translocation of Water and Solutes', 'Biochemistry and Metabolism' and 'Growth and Development'. The first two discuss how plants acquire their resources, transport them and convert them into biological materials, and translocate these to areas of growth and storage. As befits its importance, photo-

synthesis receives the greatest attention. Here the intrinsic link between structure and function is made evident. The basis of growth and differentiation is covered in the third unit. Individual chapters are devoted to the hormones or chemical regulators that influence these processes; to phytochrome, a red-light receptor that initiates numerous changes within plants; and to the control of flowering, which acquaints the newcomer with the complex interactions between external and internal factors involved in plant development.

Overall there is an impressive coverage of a wide range of topics, but inevitably some receive only cursory treatment. Figures abound in each chapter and are usually informative. The boxed essays that accompany each chapter, however, constitute a mixed bag. A few briefly describe important methods often used in physiological investigations, others focus on topics related to the text. Despite some eye-catching titles, they are largely unimpressive. Better selection of the topics could have prevented the misleading impression sometimes created by omissions in the text. It is regrettable, for example, that the wall is viewed here as a 'passive' component of the cell, and that fructans, the most important reserve carbohydrate of cereals, receive no coverage. Nevertheless, the book will please many students, not least because it deals with diverse and complex issues in a clear and manageable way. □

B. C. Jarvis is in the Department of Animal and Plant Sciences, Sheffield University, Western Bank, Sheffield S10 2TN, UK.

The war of the whorls: genetic interactions controlling flower development

Enrico S. Coen & Elliot M. Meyerowitz*

The analysis of mutations affecting flower structure has led to the identification of some of the genes that direct flower development. Cloning of these genes has allowed the formulation of molecular models of how floral meristem and organ identity may be specified, and has shown that the distantly related flowering plants *Arabidopsis thaliana* and *Antirrhinum majus* use homologous mechanisms in floral pattern formation.

MUCH of the development of flowering plants depends on meristems, the groups of dividing cells that are the source of new plant structures. On the flanks of apical meristems, found at the apex of each growing stem, additional collections of cells are set aside in defined sequences and patterns. These become either new meristems, or primordia for organs such as leaves or petals. The time and place of the formation of new meristems and organs, and their type, determine the growth and form of the plant.

In this review we discuss our current knowledge of the genetic control of meristem behaviour in flower development. In particular, we concentrate on recent advances made in the study of the meristems of two distantly related species, *Arabidopsis thaliana* (a small weed in the family *Brassicaceae*) and *Antirrhinum majus* (common snapdragon, in the *Scrophulariaceae*). In each species new mutations and newly cloned genes have revealed some of the processes that regulate meristem activity.

Arabidopsis and *Antirrhinum* development

After germination of *Arabidopsis* seed, the vegetative apical meristem produces leaves in a spiral arrangement, separated by short lengths of stem (internodes). This gives the basal part of the plant the form of a rosette. After the vegetative phase, the apical meristem reorganizes into an inflorescence meristem, which initially produces (in the same spiral pattern) a few small (cauline) leaves which will become separated by long internodes. After this transient phase, the inflorescence meristem produces floral meristems, still in a spiral pattern. Each floral meristem develops into a single flower by making concentric arrangements of the four types of floral organs in the order: sepals, petals, stamens, carpels. The result is a plant with a basal leaf rosette and an inflorescence (Fig. 1a).

Antirrhinum plants develop somewhat differently. The initial vegetative meristem produces a shoot with pairs of opposite leaves, with each pair at right angles to the previous one, and separated by long internodes. After the vegetative phase, this meristem converts into an inflorescence meristem, which produces much smaller leaves (bracts) in a spiral arrangement. The bracts are separated by short internodes, and each has a floral meristem in its axil. The floral meristems give rise to a similar series of floral organs as do the corresponding meristems of *Arabidopsis* (Fig. 1a).

The wild-type flowers of both *Arabidopsis* and *Antirrhinum* (Fig. 1b, c) consist of four whorls, or concentric regions, each occupied by organs of different types. The outer whorl (whorl 1) contains sepals; whorl 2, petals; whorl 3, stamens; and whorl 4, which occupies the centre of the flower, carpels. We will describe the identity of organs starting from the outermost whorl, so that wild-type flowers are described as sepal, petal, stamen and carpel. *Arabidopsis* flowers have four sepals, four petals,

six stamens and an ovary of two united carpels. *Antirrhinum* flowers have five sepals, five petals, four stamens (initially there are five stamen primordia but one is aborted early in development) and two united carpels (Fig. 1b). The flowers differ in symmetry: *Antirrhinum* flowers have only one plane of mirror-image symmetry (zygomorphic) whereas *Arabidopsis* flowers have two (Fig. 1c).

Homeotic changes in floral organs

Studies on the genetic control of meristem behaviour have concentrated on two classes of genes: those that control the identity of meristems, and those that determine the identity of organs. The genes in both classes can be considered homeotic, as their mutant phenotypes are the appearance of normal types of meristems or organs in positions typically occupied by other types. We begin by describing a series of homeotic mutations that affect the identity of floral organs.

Mutations are known in both species which affect the identity of organs occupying particular whorls¹⁻¹². Most of these mutations alter the identities of organs in two adjacent whorls (Fig. 2, Table 1). One class of mutants affects whorls 1 and 2, giving carpels instead of sepals in whorl 1 and stamens in place of petals in whorl 2. The overall phenotype is thus carpel, stamen, carpel (for example, *apetala2* and *ovulata* mutations have this effect). A second class of mutations affects whorls 2 and 3, and gives sepals instead of petals in whorl 2 and carpels instead of stamens in whorl 3, giving the phenotype sepal, sepal, carpel (for example, *pistillata*, *apetala3*, *deficiens*, *globosa* and *sepaloidea* mutations are in this class). Mutations in these first two classes can, depending on the locus or the particular allele, reduce organ number, in addition to affecting organ identity. A third and final class affects whorls 3 and 4 and gives petals instead of stamens in whorl 3, and sepals or variable structures in whorl 4 (*agamous* and *plena* mutations are in this class). Mutations in the third class also give extra whorls of petals or sepals inside whorl 4.

The existence of similar classes of mutations in the taxonomically distant genera *Antirrhinum* and *Arabidopsis* suggests that the mechanisms controlling floral organ identity have been highly conserved in evolution. Indeed, preliminary data show that *AP3* is homologous to *DEF* (T. Jack, L. Brockman and E.M.M., unpublished results), and *AG* to *PLE* (D. Bradley, R. Carpenter and E.S.C., unpublished results), at the level of DNA sequence as well as at the level of mutant phenotype. But there are some differences in the phenotypes of the homologous mutations in the two species, such as the appearance of sepals in the fourth whorl of *Arabidopsis ag* mutants, whereas the corresponding organs in *Antirrhinum ple* mutants are of variable type, often showing carpelloid features.

Molecular cloning of organ identity genes

Two of the organ identity genes have been cloned and extensively characterized^{13,14}. Others have been cloned more recently, with

* To whom correspondence should be addressed.

TABLE 1 Phenotype of some organ identity mutants in *Antirrhinum* and *Arabidopsis*

Genotype*	Phenotype				Region affected†
	Whorl 1	Whorl 2	Whorl 3	Whorl 4	
Wild type	Sepal	Petal	Stamen	Carpel	
<i>ovu, ap2</i>	Carpel	Stamen	Stamen	Carpel	A
<i>def, glo, sep, pi, ap3</i>	Sepal	Sepal	Carpel	Carpel	B
<i>plena, ag</i>	Sepal	Petal	Petal	Variable‡	C

* The *Arabidopsis* mutations are *apetala2* (*ap2*), *pistillata* (*pi*), *apetala3* (*ap3*) and *agamous* (*ag*). The *Antirrhinum* mutations are *ovulata* (*ovu*), *deficiens* (*def*), *sepaloides* (*sep*), *globosa* (*glo*) and *plena* (*ple*). Several mutations at a locus called *pleniflora* have recently been described in *Antirrhinum*⁸; complementation tests have shown that these are allelic to the classic *plena* mutation²² (R. Carpenter and E.S.C., unpublished results).

† See Fig. 3.

‡ In *Antirrhinum*, whorl 4 can be petaloid, sepaloid, carpelloid or a mixture of these; in *Arabidopsis* this whorl contains sepals. In both species additional petaloid and sepaloid whorls are produced interior to whorl 4.

initial characterization still in progress (ref. 9; T. Jack, L. Brockman and E.M.M., manuscript in preparation; D. Bradley, R. Carpenter and E.S.C., unpublished results; R. Simon, R. Carpenter, S. Doyle and E.S.C., unpublished results). The first homeotic organ identity genes to be cloned were *DEFICIENS* of *Antirrhinum*, with a class two phenotype of sepal, sepal, carpel, carpel¹³, and *AGAMOUS* of *Arabidopsis*, with the class three phenotype sepal, petal, petal, sepal, followed by additional whorls of sepals and petals¹⁴. Remarkably, the proteins encoded by each gene show extensive similarity. Each contains a region of 56 amino acids, of which 40 are identical in the two proteins. There is thus 71% amino-acid identity between these parts of the proteins. The highly conserved region in these genes is very probably involved in DNA binding; it shows considerable similarity to regions in the vertebrate serum response factor genes and to the yeast *MCM1* gene. Serum response factors are DNA-binding transcriptional regulators of the *c-fos* oncogene in humans, and of actin genes in *Xenopus*^{15,16}. The yeast *MCM1* gene encodes PRTF, a DNA-binding transcriptional regulator of mating type-specific genes^{17,18}. The functional motif shared by all these genes has recently been named the MADS box (*MCM1-AGAMOUS-DEFICIENS-SRF*) and it has been shown that additional plant homeotic genes also contain a MADS box^{9,19}. Thus, some of the homeotic genes of plants, like those of insects, are members of a family of DNA-binding proteins, each member serving different functions in the organism.

There is an additional region in the centre of both the AG and DEF proteins that shows structural similarity: in each protein, this region (called the K-box because of its similarity to the region of keratin responsible for coiled-coil formation) could form two alpha helices, highly charged on one face, and mostly uncharged on the other¹⁹. The function of this region is unknown, though one could imagine a role in protein-protein interactions.

Model of control of organ identity

We define three regions of the floral meristem, each coincident with the domain of action of one of the three classes of floral homeotic genes. Region A comprises whorls 1 and 2; region B, 2 and 3; and region C, 3 and 4 (Fig. 3). The action of several genes in these overlapping regions could give each whorl a unique combination of functions⁵. For example, if genes acting in regions A, B and C are required for three regulatory functions *a*, *b* and *c*, respectively, then the combination of functions in the four whorls of wild-type would be: *a*, *ab*, *bc*, *c* (Fig. 4). In principle, this might provide sufficient information to specify the identity of organs in each whorl. That is, sepals form if *a* alone is expressed, *a* and *b* together direct petal development, *b* and *c* together specify stamens, and *c* expressed alone determines carpel formation.

The homeotic mutations described above each eliminate one of the postulated functions: the *ap2* and *ovu* mutations disrupt

function *a*, *ap3*, *pi*, *def*, *glo* and *sep* mutations prevent function *b*, and mutations in *ag* and *ple* block function *c*. An important constraint on such a combinatorial model is that it needs to account for the mutant phenotypes observed when the genes necessary for the *a*, *b* or *c* functions are mutated. By this criterion, the simple model that the domains of *a*, *b* and *c* function are established independently of each other does not fit the data. For example, if *a* is required for sepal and petal development in region A, how can these organ types develop outside this region in certain mutants, as for example in the *ag* and *ple* mutants, which lack the *c* function, and have petals in whorl 3?

To account for this effect, it has been proposed that *a* and *c* might influence each other's expression^{8,10,12}. Strong evidence for an interaction between *a* and *c* has come from the study of double and triple mutants (Fig. 5). Triple mutants lacking *a*, *b* and *c* have been constructed in *Arabidopsis* and give flowers consisting solely of organs resembling cauline leaves^{10,12} (Fig. 5d). These leaves can be considered as a type of developmental 'ground state'. Starting from this state we now consider the phenotypes observed when the various functions are added back (Fig. 4). Addition of *a* alone to the ground state results in a flower with only sepals, as observed in *bc* double mutants in both *Arabidopsis* and *Antirrhinum* (refs 5, 10; R. Carpenter and E.S.C., unpublished results). Addition of *c* to the ground state gives only carpels, as seen in *ab* double mutants (refs 5, 10; R. Carpenter and E.S.C., unpublished results). But addition of both *a* and *c* results in the phenotype sepal, sepal, carpel, carpel (that is, *b* mutants); showing that the actions of *a* and *c* are now restricted to the respective regions A and C. Thus, the *a* and *c* functions seem to be antagonistic, and establish mutually exclusive domains of action.

One notable difference between the *a* function mutations in the two species is that they are recessive in *Arabidopsis*, and semidominant in *Antirrhinum*. This might indicate that in *Antirrhinum* the mutations inhibit *a* function rather than cause loss of this function. One possibility is that the dominant *a* mutations are in fact gain-of-function alleles of *c*, and that the gene providing the *a* function is not yet genetically identified in *Antirrhinum*.

The domain of *b* function is established independently of the *a* and *c* functions. This is demonstrated by the phenotype of *Arabidopsis* flowers with only the *b* activity (that is, *a*, *c* double mutants), which have leaflike ground state organs in whorls 1 and 4, and organs intermediate between petals and stamens in whorls 2 and 3 (refs 5, 10; Fig. 5b).

The *a*-*b*-*c* model for specification of organ identity predicts that the activity of the *b* genes is in region B, and that of the *c* genes in region C. This spatial regulation of gene activity seems, in this case, to be at the RNA level, as the *Antirrhinum* DEF and *Arabidopsis* AP3 gene RNAs are found by *in situ* hybridization to be predominantly in petals and stamens and their primordia (region B), and the *Arabidopsis* AG gene RNA predominantly in stamens and ovary and their primordia

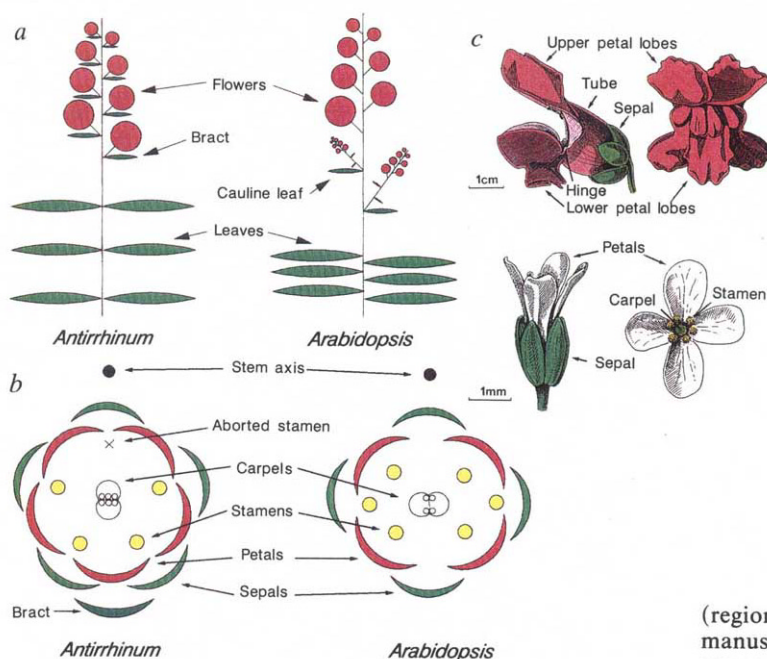


FIG. 1 *a*, Schematic diagrams of wild-type *Antirrhinum majus* and *Arabidopsis thaliana* plants. Lateral shoots arising in the axils of leaves are omitted for clarity. *b*, Floral diagrams of wild-type *Antirrhinum* and *Arabidopsis*. The small circles in the ovaries represent ovules. *c*, Flowers of *Antirrhinum* (above) and *Arabidopsis* (below). Flowers are shown in side view (left) or face view (right). The *Antirrhinum* flower shown in side view is opened slightly to illustrate the hinge between upper and lower petals. The face view of *Arabidopsis* is at a 45° orientation relative to the floral diagram in *b*. In linear dimension, *Arabidopsis* flowers are close to 10 times smaller than those of *Antirrhinum*, and in mass they are more than 1,000 times smaller. Nonetheless, the basic processes of development, and their results, are similar. *Antirrhinum* drawings adapted from ref. 27; *Arabidopsis* drawings by K. Roberts.

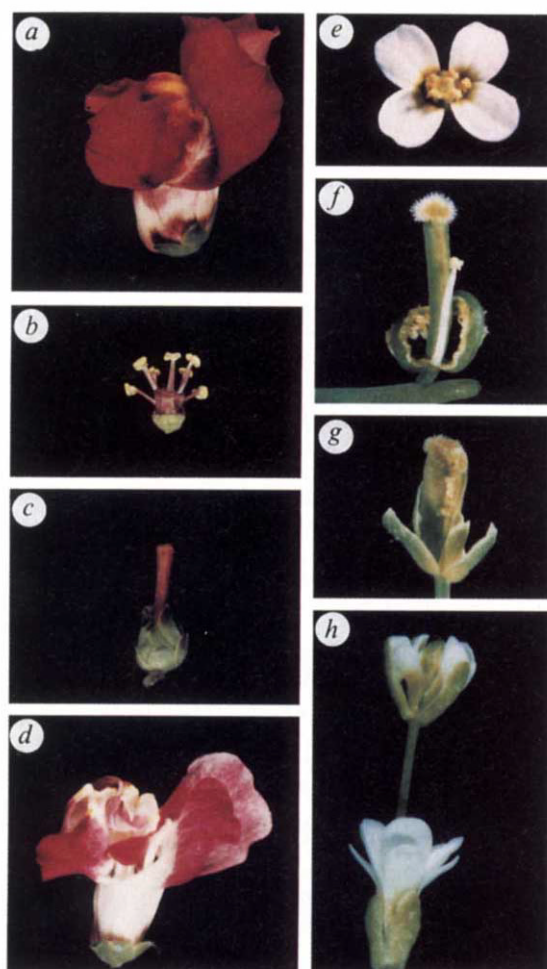


FIG. 2 Photographs of the parallel series of homeotic mutants of *Antirrhinum* (left) and *Arabidopsis* (right). *a* and *e*, Wild-type. *b* and *f*, The class one mutants *ovulata* and *apetala2* (the depicted alleles are *ovu-621* and *ap2-2*). *c* and *g*, The class two mutants *deficiens-621* and *pistillata-1*. *d* and *h*, The class three mutants *plena-624* and *agamous-2*. The *agamous* flower (*h*) shows clearly the development of the flower whorls interior to whorl 4, owing to the elongation of the pedicel of the first inner flower.

(region C) (refs 9, 13, 14, 20; T. Jack, L. Brockman and E.M.M., manuscript in preparation). This pattern is consistent with the proposed model, and also provides the opportunity to test a key element in the model, the *a-c* antagonistic interaction. In *Arabidopsis*, if the spatial control of *AG* expression is at the RNA level, then the domain of expression of *AG* RNA should expand to all four whorls in an *apetala2* background (Fig. 4). This has recently been shown to be true²⁰. The molecular evidence thus accords well with the predictions of the genetic model for organ identity.

The *a-b-c* model raises the question of how the organ identity genes are themselves regulated as, for example, the *b* function genes act only in region B, regardless of the activity of the other homeotic genes. There is an *Arabidopsis* gene (*SUPERMAN*) whose mutant phenotype is consistent with a role in regulation of the spatial pattern of the *b* function genes^{10,12}. Recessive mutations in it reduce or eliminate the fourth whorl carpels, and increase the number of stamens, which may occupy the central region of the flower (Fig. 6a). The wild-type gene thus appears to repress the *b* function in the fourth whorl. Additional gene functions must also be necessary for proper expression of the organ homeotic genes.

Genetic control of whorl and organ number

Mutations in genes needed for the *c* function generally give an increase in whorl number and an indeterminate growth pattern. For the *ag* mutants in *Arabidopsis*, the mutant phenotype is sepal, petal, sepal, petal, sepal, petal, sepal and so on. This is sometimes described as 'flowers within flowers' because the sepals in whorls 4, 7 and so on can also be considered as the first whorl of sepals of an enclosed flower¹⁰. The *ple* mutants in *Antirrhinum* are similar to *ag* mutants, but give variable organs in whorl 4. These findings indicate that the *c* organ identity genes have two roles: the control of organ fate, and activation of a determinacy function to prevent indeterminate floral meristem growth.

Some of the other organ identity mutants also affect organ or whorl number. In one case this seems to be a result of the ectopic action of the determinacy function of the *c* genes. Extreme mutants of *AP2*, a gene required for the *a* function in *Arabidopsis*, give fewer organs than wild type in whorls 1, 2 and 3. As determinacy can be caused by the *c* function in whorl 4, and as *a* mutants have ectopic activity of the *c* function in outer whorls, this could explain the reduced organ numbers. Evidence favouring this interpretation comes from counting organ number in *Arabidopsis* mutants lacking *a* or lacking both *a* and *c*. The first three whorls have, in wild-type flowers, a total of 14 organs. An

extreme mutant lacking *a* (*ap2-2*) can have as few as two organs in these whorls. The removal of *c* function from such flowers, by making the doubly mutant *a c* (*ap2-2 ag*) strain, restores most of the missing organs¹⁰.

It should be noted that the gene required for *a* function is not necessary for proper organ identity in the third whorl, but that it is necessary (at least in *Arabidopsis*) for proper organ number in that whorl. The mechanisms of organ identity specification and establishment of organ number are thus to some degree separate, even though some of the same gene products are involved. Another indication that establishment of organ number and organ identity may be separate processes is the existence of mutations that change organ number without affecting organ identity; for example *clavata1* of *Arabidopsis*, which causes an increase in the numbers of organs in whorls 2, 3 and 4 (ref. 2).

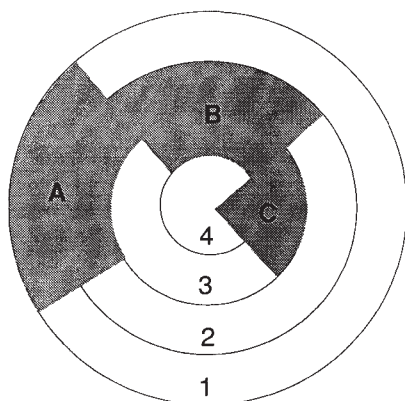


FIG. 3 Schematic diagram of the four whorl regions of a floral primordium (viewed from above, as in Fig. 1b), and the three regions, A, B and C, which are the domains of action of the three classes of organ homeotic genes.

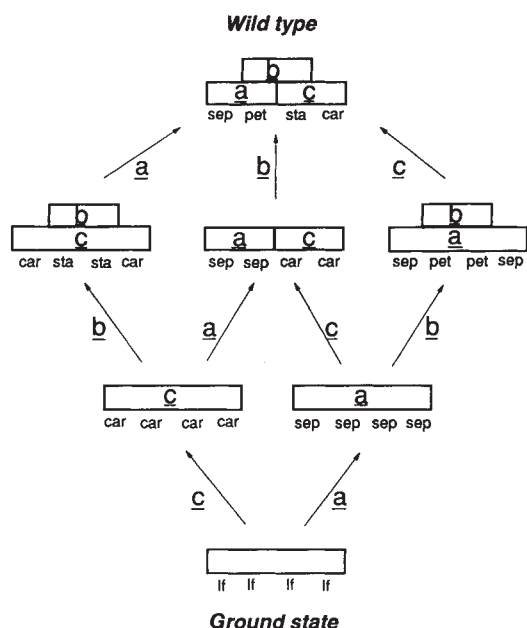


FIG. 4 Model illustrating the combinatorial interaction of the *a*, *b* and *c* functions in wild-type and various single, double and triple mutant combinations. The phenotype of the triple mutant is shown (and labelled ground state) at the bottom. The phenotypes observed when different combinations of functions are 'added back' to the ground state are shown above. If, Leaf with some carpelloid features; car, carpel; sep, sepal; pet, petal; sta, stamen.

Genetic control of differences in whorls

In *Arabidopsis*, all organs in the same whorl have similar morphologies. In *Antirrhinum* and other species whose flowers are zygomorphic, one or more whorls contain organs with distinct morphologies. For example, in whorl 2 the upper two petal lobes of *Antirrhinum* have a distinct shape from the lower three, and in whorl 3 the uppermost stamen is aborted early in development (Fig. 1). Many mutations have been described in zygomorphic species that can reduce or eliminate the differences between organs in a whorl and thus render the flower more or less radially symmetrical^{6,11}. These represent a separate class of organ homeotic genes from those discussed above, because they change organs to a type normally found in the same whorl, and not in a different one. In these mutants all of the organs in a whorl resemble one particular organ (generally the upper or lower organ) of the corresponding wild-type whorl.

The most intensively studied example of this phenomenon is in *Antirrhinum*. Several different mutations of the *cycloidea* (*cyc*) gene have been described which give flowers with a more symmetrical appearance than wild type^{8,21,22} (Fig. 7). Extreme *cyc* mutations give radially symmetrical flowers with all petals resembling the lowest petal of wild type. Unlike wild type in which the uppermost stamen is aborted, all five stamen primordia develop fully in these mutants to give mature organs of similar length to the lower stamens of wild type. Thus, all organs in whorls 2 and 3 resemble the lowest organs of the corresponding whorl in the wild-type flower. In addition to this phenotype, there are also many *cyc* alleles which confer intermediate phenotypes, ranging from almost symmetrical flowers to nearly wild type.

As described earlier, many mutations give organs with identities inappropriate to the whorl in which they develop, so it is possible to ask if the action of *CYC* on particular organs depends on the whorl which they occupy. For example, the uppermost organ in whorl 3 of wild-type *Antirrhinum* is aborted as a result of *CYC* action because in extreme *cyc* mutants all five stamen primordia develop fully in whorl 3 (Fig. 1). If stamens grow in whorl 2, as in *ovu* mutants, the upper two organs of this whorl are also vestigial or aborted. The two upper organs are therefore aborted whether stamens grow in whorls 2 or 3. Furthermore, in *cyc ovu* double mutants, all stamen primordia develop fully in whorls 2 and 3 to give a flower with 10 stamens (R. Carpenter and E.S.C., unpublished results). This suggests that *CYC* interacts with primordia in a similar way, irrespective of the whorl that they occupy, and therefore that the fate of a primordium depends on an interaction between functions determining whorl identity and those determining the differences between upper and lower organs. These observations have suggested a polar-coordinate model for the control of primordium fate in each whorl⁸ (Fig. 8).

Homeotic changes in floral meristems

Some mutations affecting earlier regulatory steps than those under the control of the organ homeotic genes might be expected to prevent or alter the formation of floral meristems. Mutants of this type, in which inflorescence meristems (or structures intermediate between inflorescence meristems and floral meristems) appear in the positions normally occupied by floral meristems, are *floricaula* (*flo*) and *squamosa* (*squa*) in *Antirrhinum*^{8,9} and *leafy* (*lfy*) and *apetala1* (*ap1*) in *Arabidopsis*^{3,23}. In *flo* mutants of *Antirrhinum* vegetative growth and the transition to the inflorescence meristem is similar to wild type. Instead of floral meristems being produced in the axils of bracts, however, the meristems in these positions are of the inflorescence type (Fig. 9). Thus, indeterminate shoots are produced in the bract axils, and each of these shoots can in turn produce further shoots in the axils of their bracts. *Arabidopsis lfy* and *ap1* and *Antirrhinum squa* mutants have a similar phenotype, though they show only partial conversion of floral meristems to inflorescence meristems (Fig. 6b).



FIG. 5 Double and triple mutant strains of *Arabidopsis*. *a*, Flowers lacking both the *b* and *c* functions (genotype *ap3-1 ag-1*), and as a consequence having sepals as the only floral organs. *b*, Flowers lacking both the *a* and *c* functions (genotype *ap2-1 ag-1*). This causes the organs of whorls 1 and 4 to be leaflike, whereas the organs of whorls 2 and 3 are intermediate between petals and stamens. *c*, Flowers lacking the *a* and *b* functions (genotype *ap2-2 ap3-1*). All organs are carpels, and as the 'strong' *ap2-2* allele of *apetala2* was used, many organ positions are unoccupied. *d*, Flowers lacking all three organ homeotic gene functions, and containing only leaflike ('developmental ground state') organs in the positions normally occupied by floral organs (genotype *ap2-1 ap3-1 ag-1*). The leaflike organs have some carpelloid features such as stigmatic tissue at the tips.

The *FLO* gene of *Antirrhinum* has been cloned and characterized²⁴. It produces a transcript which has the potential to encode a protein of 396 amino acids. The protein has a proline-rich amino terminus and an acidic region, both features of transcriptional activators^{25,26}. But the sequence shows no extensive similarity with the organ specification genes or other sequences available in data banks, so that a role other than transcriptional regulation is not excluded. *In situ* hybridization shows that *FLO* is expressed very early in wild-type inflorescences. Earliest expression is in bract primordia and is followed by expression in sepal, petal and carpel primordia, but no expression is seen in stamen primordia. Expression in each organ is transient and is not observed in later stages of development. Taken together, these results suggest that *FLO* protein not only acts as a switch between inflorescence and floral meristems, but also may be involved in directing or maintaining specific patterns of gene expression in the early floral meristem. One can speculate that the expression of *FLO* in certain primordia may activate genes, such as organ identity genes, required for their normal development. Similarly, the absence of *FLO* from whorl 3 may be necessary for normal stamen development²⁴. If true, an additional piece of the floral development puzzle will begin to fall into place: the spatial patterns of expression of the organ homeotic genes are the key to organ identity in developing flowers, but these genes seem to be regulated by an earlier-acting group of genes, some of which have meristem-conversion phenotypes.

The pattern of *FLO* expression also illustrates that early and late genes need not interact in a simple hierarchical fashion. For example, the absence of *FLO* RNA in whorl 3 might be explained if the combination of organ identity gene functions in this whorl (*b* and *c*) inhibited *FLO* expression²⁴. This hypothesis has been recently confirmed by the observation that *FLO* RNA is present in whorl 3 of mutants lacking *b* or *c* (S. Hantke, R. Carpenter and E.S.C., unpublished results). Thus,

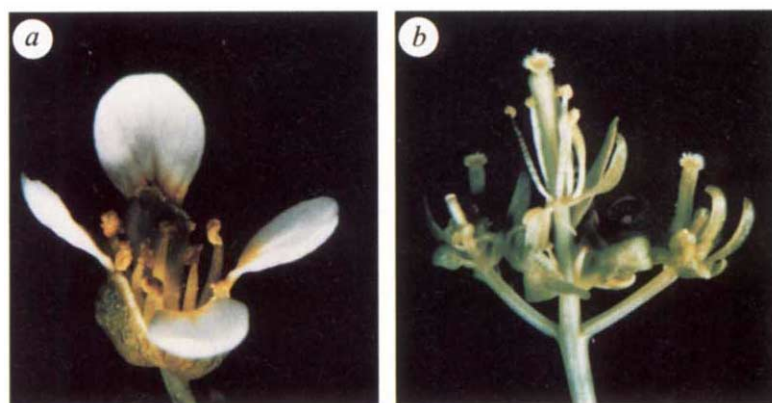


FIG. 6 Additional *Arabidopsis* mutations, thought to interact with the organ identity genes. *a*, A *superman-1* homozygote, showing extra stamens and reduced ovary. One wild-type function of this gene seems to be preventing *b* function in whorl 4. *b*, An *apetala1-1* homozygote, showing partial conversion of floral primordia to inflorescence meristems. Note that each pedicel is capped by a group of flowers, rather than a single flower.



FIG. 7 Face view of wild-type (left) and extreme *cycloidea* (right) *Antirrhinum* flowers. In *cyc* mutants all petals resemble the lowest petal of wild type. Because the lowest petal-lobe of wild type tends to fold back against the tube (Fig. 1), in *cyc* mutants all of the petals are folded back. The genetic backgrounds of the lines illustrated carry mutations in two pigmentation genes, resulting in orange petals.

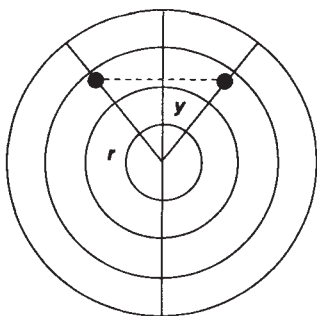


FIG. 8 Polar coordinate model for bilateral symmetry in *Antirrhinum*. The four whorls are shown as concentric rings. The vertical line (y axis) indicates where the plane of symmetry bisects the flower (the symmetry plane is at right angles to the plane of the page). The identity of organs at each position in the flower can be determined by direct comparison with Fig. 1*b*. Upper organs are towards the top of the y axis and lower organs towards the bottom. Expression of organ identity functions varies along the radius (r) and the *cyc* function varies along the y axis, increasing in activity towards the top of the flower. Two primordia with the same combination of functions and, hence, the same developmental fate, are shown joined by a dotted horizontal line (these correspond to the two upper petal primordia, see Fig. 1*b*). Mutations eliminating the organ identity functions result in some primordia in different whorls having similar specifications so that they adopt similar developmental fates. Mutations that abolish *cyc* function remove differential expression along the y axis such that all primordia in a whorl adopt a fate similar to that of the lower primordia of the wild-type whorl.

even though *FLO* is activated earliest in the floral meristem as a whole, in cells giving rise to whorl 3 organs, *b* and *c* functions could be established earlier than *FLO* and therefore regulate its expression.

Timing of gene action

It is therefore possible that the time and duration of gene action, as well as its place, is of central importance to understanding the genetic interactions underlying floral development. The timing of *DEF* gene function in *Antirrhinum* and *AP2* and *AP3* gene function in *Arabidopsis* have been studied. One of the effects of *def* mutations is that sepals grow in place of petals in whorl 2, presumably because the *b* function is absent. The unstable allele, *def-621*, shows distinct clonal patches of petal tissue on the sepals in whorl 2 (ref. 8). These patches are separated from the surrounding sepal epidermal cells and the underlying mesophyll tissue by sharp boundaries. They can be explained by somatic excision of a transposon from the *DEF* locus, restoring gene function and hence petal morphology. Some of the patches consist of as few as four cells, presumably reflecting restoration of *DEF* function in the last few cell divisions of organ development. This indicates that *DEF* can act at late stages to direct cells of whorl 2 toward a petal developmental program. But the phenotypic effects of the *def* mutation can also be detected at early developmental stages, when petal and sepal primordia acquire distinct morphologies. The *DEF* product is thus active in whorl 2 throughout its development, from the early stages when petal primordia become distinct to the final cell divisions of the petal. This is consistent with the observed expression pattern of *DEF* throughout flower development¹³.

Additional evidence for the action of genes required for the *b* function over an extended period has come from studies on the effects of changing environmental conditions during flower morphogenesis. One *ap3* allele in *Arabidopsis* confers a temperature-sensitive phenotype, with a restrictive temperature of 29 °C, and a permissive one of 16 °C. By shifting the growth temperature during organogenesis, it has been shown that *AP3* activity is required in whorl 2 from early periods up to relatively late stages of flower development, including the time when visible differentiation is in progress⁵. As *ap3* seems to affect the

same *b* function as does *def*, these results are consistent with the mosaic studies.

In addition to genes acting for extended periods of development, there are other genes with a more transient action. For example, temperature-shift experiments with certain *ap2* alleles (required for the *a* function) indicate that *AP2* is not required at any time after an early stage of flower development⁵. In this respect, the action of class *a* genes might be quite different from that of class *b*, at least in the second whorl. It should be recognized, however, that the temperature-sensitive period of these *ap2* alleles may be a function of the particular allele, or of an underlying temperature-sensitive part of flower development, and may therefore not indicate the full period of *AP2* activity.

Conclusion

Plant form is largely a result of meristem behaviour. We have distinguished between meristems, which produce other meristems or collections of organs, and organ primordia, which develop to a single recognizable organ. We have described on this basis two classes of homeotic mutation: those which affect meristem identity, and those which affect the fate of floral organ primordia. This last class contains two types: organ identity can be inappropriate to a floral whorl, or in the case of zygomorphic flowers, inappropriate to a position in a whorl. Study of the phenotypes of the mutations, of the patterns of expression of the RNAs coded by the genes, and of the predicted amino-acid sequences of the gene products has led to specific models for the way in which these genes direct flower development. Further study of these and other mutations may in the future lead toward solutions to additional problems in flower development, such as the way in which a vegetative meristem converts to an inflorescence meristem, or the nature and control of the complex patterns of cell differentiation necessary in each organ after its identity has been specified.

Two general conclusions can be reached from the comparison of flower development genes in two distantly related species. The first is that the basic mechanisms that define organ identity in developing flowers appear to be the same in both species. Genes exist whose mutants have similar phenotypes and similar interactions, and a single model can explain the action of these genes in both species. In those cases known so far, the genes with similar phenotypes are homologous at the DNA level. This



FIG. 9 The floricaula homozygous *Antirrhinum* inflorescence, in which each flower is replaced by a shoot.

shows that the basic processes of floral organ specification are evolutionarily old, and that the very different forms taken by flowers of different plant families, at least among the dicotyledonous plants, result from more recent modifications of developmental processes. This recognition allows experimental approaches to morphological evolution as wild-type or altered *Antirrhinum* genes can now be introduced into *Arabidopsis* plants with mutations in the homologous gene. When *Antirrhinum* transformation procedures are discovered, the converse experiment will also be possible. Such experiments will reveal whether the evolutionary changes that lead to the very different flower forms of the two plants are due to changes in the homeotic genes themselves, changes in their regulators, or changes in the downstream genes that they regulate.

The second conclusion is that the processes of flower development that have been revealed are remarkably size-invariant. *Arabidopsis thaliana* flowers, when mature, have a mass of around 550 µg, whereas *Antirrhinum majus* flowers have a mass more than 1,000 times greater. Nonetheless, the basic processes that specify their meristem and organ types appear to be homologous. We soon may learn how the relative timing of cell division and regulatory gene expression are controlled, and how the domains of gene expression are developmentally regulated, and thus comprehend how similar structures of very different sizes may evolve from a common ancestor. □

Enrico S. Coen is at the Department of Genetics, John Innes Institute, Colney Lane, Norwich NR4 7UH, UK; Elliot M. Meyerowitz is at the Division of Biology, California Institute of Technology, Pasadena, California 91125, USA.

- Pruitt, R. E., Chang, C., Pang, P. P.-Y. & Meyerowitz, E. M. in *Genetic Regulation of Development* (ed. W. R. Loomis) 327-338 (Liss, New York, 1987).
- Bowman, J. L., Yanofsky, M. F. & Meyerowitz, E. M. *Oxford Surv. Plant molec. cell. Biol.* **5**, 57-87 (1988).
- Haughn, G. W. & Somerville, C. R. *Dev. Genet.* **9**, 73-89 (1988).
- Komaki, M. K., Okada, K., Nishino, E. & Shimura, Y. *Development* **104**, 195-203 (1988).
- Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. *Pl. Cell* **1**, 37-52 (1989).
- Meyerowitz, E. M., Smyth, D. R. & Bowman, J. L. *Development* **106**, 209-217 (1989).
- Kunst, L., Klenz, J. E., Martinez-Zapater, J. & Haughn, G. W. *Pl. Cell* **1**, 1195-1208 (1989).
- Carpenter, R. & Coen, E. S. *Genes Dev.* **4**, 1483-1493 (1990).
- Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. & Sommer, H. *Science* **250**, 931-936 (1990).
- Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. *Development* **112**, 1-20 (1991).
- Coen, E. S. *A. Rev. Pl. Physiol. Pl. molec. Biol.* **42**, 241-279 (1991).
- Meyerowitz, E. M. *et al. Development* **112** suppl. 1, 157-168 (1991).
- Sommer, H. *et al. EMBO J.* **9**, 605-613 (1990).
- Yanofsky, M. F. *et al. Nature* **346**, 35-39 (1990).
- Norman, C., Runswick, M., Pollock, R. & Treisman, R. *Cell* **55**, 989-1003 (1988).
- Boxer, L. M., Prywes, R., Roeder, R. G. & Kedes, L. *Molec. cell. Biol.* **9**, 515-522 (1989).
- Passmore, S., Maine, G. T., Elble, R., Christ, C. & Tye, B. K. *J. molec. Biol.* **204**, 593-606 (1988).

- Herskowitz, I. *Nature* **342**, 749-757 (1989).
- Ma, H., Yanofsky, M. F. & Meyerowitz, E. M. *Genes Dev.* **5**, 484-495 (1991).
- Drews, G. N., Bowman, J. L. & Meyerowitz, E. M. *Cell* **65**, 991-1002 (1991).
- Kuckuck, H. Z. *Indukt. Abst. Vererbungsl.* **56**, 51-83 (1936).
- Stubbe, H. *Genetik und Zytologie von Antirrhinum L. sect. Antirrhinum* (Gustav Fischer, Jena, 1966).
- Irish, V. F. & Sussex, I. M. *Pl. Cell* **2**, 741-753 (1990).
- Coen, E. S. *et al. Cell* **63**, 1311-1322 (1990).
- Mermod, N., O'Neill, E. A., Kelly, T. J. & Tjian, R. *Cell* **58**, 741-753 (1989).
- Struhl, K. *Trends biochem. Sci.* **14**, 137-140 (1989).
- Weberling, F. *Morphology of Flowers and Inflorescences* (Cambridge University Press, Cambridge, 1989).

ACKNOWLEDGEMENTS. We thank the members of our laboratories for reading and criticizing this manuscript, and J. Bowman for help with the figures. The work of the Meyerowitz laboratory has been supported by the US National Science Foundation (Genetics), the US National Institutes of Health (General Medicine), the US Department of Energy (Energy Biosciences) and the Markey Foundation. The Coen laboratory is supported by the Gatsby Foundation, the European Community, the US National Science Foundation, and the Agricultural and Food Research Council. Our collaborative work is funded by the Human Frontier Science Program Organization.

Tomographic imaging of subducted lithosphere below northwest Pacific island arcs

Rob van der Hilst*, Robert Engdahl†, Wim Spakman‡ & Guust Nolet‡§

* Department of Earth Sciences, University of Leeds, Leeds LS2 9JT, UK

† National Earthquake Information Center, US Geological Survey, DFC, Box 25046, Stop 967, Denver, Colorado 80225, USA

‡ Department of Theoretical Geophysics, University of Utrecht, Budapestlaan 4, 3508 TA Utrecht, The Netherlands

The seismic tomography problem does not have a unique solution, and published tomographic images have been equivocal with regard to the deep structure of subducting slabs. An improved tomographic method, using a more realistic background Earth model and surface-reflected as well as direct seismic phases, shows that slabs beneath the Japan and Izu Bonin island arcs are deflected at the boundary between upper and lower mantle, whereas those beneath the northern Kuril and Mariana arcs sink into the lower mantle.

near the transition between the upper and lower mantle to learn more about the amount of mass transport across this boundary¹. Although many seismologists explain the lower-mantle heterogeneity beneath Middle America on the basis of deep subduction of the Farallon/Pacific plate²⁻⁸, the subject has remained controversial in studies of northwest Pacific subduction zones.

In the northwest Pacific, where old lithosphere of the Pacific plate subducts below island arcs⁹, Jordan and coworkers¹⁰⁻¹³ concluded from residual sphere analyses that the subducted slab continues below the deepest earthquakes, to depths of at least 1,200 km. There is some evidence to support this interpretation¹⁴⁻¹⁷; other studies, however, have suggested that the images of subducted slabs determined by this method can be influenced by noise in the data and by effects of lower-mantle and near-receiver structure¹⁸⁻²².

Variations in the propagation velocity of seismic waves owing to the presence of subducted lithosphere can be imaged by seismic tomography. The tomographic method employed in this study involves the interpretation of seismic-wave arrival-times determined from millions of seismograms in terms of the Earth's three-dimensional structure. The controversial issues concerning deep subduction have not so far been satisfactorily resolved by

THE depth range involved in the recycling of material from the Earth's surface back into the mantle has been among the most controversial issues in geodynamics in the past two decades. Many studies have investigated the fate of subducted lithosphere

§ Now at: Department of Geological and Geophysical Sciences, Princeton University, Princeton, NJ 08544, USA.

tomographic studies, mainly because of the lack of high-quality seismic data. Another problem is the strong nonlinearity of the tomographic problem, which requires linearization with respect to a background Earth model that closely mimics the real Earth. The solutions of linearized inversions are necessarily close to this *a priori* model. Published results are equivocal but mathematically acceptable as 'minimum norm solutions'²³ of the inversion, and resolution tests have been used to claim their reliability. The tests commonly used to assess the reliability of tomographic images determined by linear inversion do not, however, address the influence of the *a priori* assumptions and data quality.

Our objectives are to demonstrate that the use of different seismic data and *a priori* assumptions can easily result in different tomographic images, and that improvements in the reference Earth model and, as a consequence, in the seismic data and earthquake hypocentres routinely reported by international data centres, result in better tomographic images. Important innovations to current methods introduced in this study are the recomputation of earthquake hypocentres and the re-identification of associated P- and pP-wave data using the reference model iasp91 (ref. 24) before the inversion for three-dimensional structure.

The iasp91 model is a good description of the spherically averaged velocities of longitudinal waves below the study region, which is a requirement for a proper linearization of the inversion problem. The use of the iasp91 model and data from the direct P and surface-reflected pP waves from improved hypocentres provides information about deep slab structure that is more reliable than previous inversions based solely on P data and hypocentres reported by the International Seismological Centre (ISC). The images resulting from our improved method reveal relatively short slabs with high seismic velocity relative to surrounding mantle material. In contrast, the inversion of uncorrected ISC P-wave data resulted in long slabs, penetrating to lower-mantle depths, with smaller variations in seismic velocity. The new images indicate slab deflection at the 670-km discontinuity from southern Kuril to central Izu Bonin, but suggest mass transport to lower-mantle depths below the northern Kurile and the Mariana arcs.

Advantages of pP data

In the northwest Pacific, earthquakes and seismological stations are confined mainly to narrow zones along plate boundaries (Fig. 1). When the data set consists entirely of direct P-wave data, this distribution of earthquakes and stations causes a number of problems in the resolution of subduction-zone structure by tomographic imaging²⁵. First, shallow structures below intraplate regions will not be recovered because of inadequate sampling. Second, because direct waves travel predominantly either upwards to stations along the arc or steeply downwards

to more-distant stations, and only few waves are oblique to the slab plane, 'smearing' of structure along the ray paths will occur and slab structure cannot be resolved unambiguously²⁶. Third, an uneven station distribution can result in biased or poorly constrained hypocentres and loss of structurally related signal in the data.

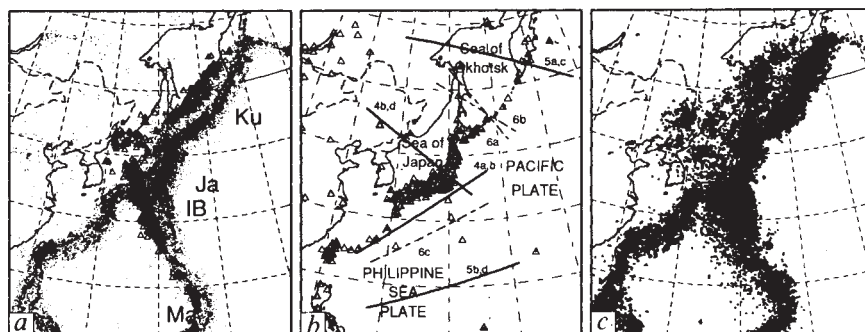
These problems are reduced by augmenting the data set used for the tomographic imaging by data from the seismic phase pP (ref. 25). This phase is a longitudinal wave that initially radiates upwards from the hypocentre, is reflected at the Earth's surface, and subsequently travels to a distant seismological station. The pP phase is sometimes called a depth phase because the phase can be observed only for earthquake hypocentres with non-zero depth. The northwest Pacific region is very suitable for the incorporation of the pP phase in tomographic imaging because of the numerous deep earthquakes and, in most cases, the substantial spatial separation between P- and pP-wave ray paths. The distribution of pP reflection points (Fig. 1c) indicates that pP-wave ray paths geometrically sample mantle structure beneath the intraplate regions more adequately than do direct P waves. Moreover, pP-wave ray paths are often oblique to ray paths of the direct P waves. This enhances the resolution along P-wave ray paths and thus counteracts the 'smearing' of velocity anomalies. In addition, the combined use of P- and pP-wave arrival times puts strong constraints on the earthquake focal depth^{27,28} which limits the trade-off between uncertainties in earthquake location and variations in P-wave velocity.

In spite of these obvious advantages, the pP phase has not been used in tomographic imaging until recently because the data were suspected of being unreliable. Van der Hilst and Engdahl²⁵, however, investigated possible sources of error and showed that, after careful processing, pP data could significantly improve tomographic images of mantle structure beneath the Caribbean region. The statistical properties of pP data reported by the ISC for northwest Pacific earthquakes are summarized in Fig. 2a. The frequency distribution of the ISC pP data is skewed to negative values but has a large positive tail. This tail is due mainly to misidentifications of other later-arriving depth phases such as sP (initially transverse compressional wave) and water-reflected pwP, most of which can be recognized and removed before inversion (R.E. and R.v.d.H., manuscript in preparation, hereafter referred to as EH91).

Earthquake relocation

In delay-time tomography, *a priori* knowledge of seismic velocities in the Earth's interior (the reference model) is used to linearize the inversion problem²³. This reference model is then used for estimating earthquake locations, for calculating travel times, and for computing approximate ray paths along which the seismic waves travel from source to receiver. The delay time, or travel-time residual, is the difference between the

FIG. 1 a, Epicentres of the northwest Pacific earthquakes used in this study (time period: 1964–1989). Two intervals of focal depth are distinguished: dots, 0–300 km ($N=38,484$) and triangles, 300–700 km ($N=1,753$). Abbreviations denote the Kuril (Ku), Japan (Ja), Izu Bonin (IB) and Mariana (Ma) island arcs. b, Locations of the seismological stations within the study region. The 323 stations contributed 544,865 travel-time residuals to the data set. We also used data from 1,961 stations outside the study region, primarily in Europe and North America, which contributed 1,363,431 P-wave data. These data are corrected for delays acquired along the P- and pP-wave ray paths outside the Earth's volume under study (see text for discussion). The solid and dashed lines depict the strikes of the mantle cross-sections shown in Figs 4 and 5 and Fig. 6 respectively. c, The distribution of 83,831 pP reflection points. The reflection points are computed in a spherical Earth



model. The pP data (Fig. 2c) are corrected for water depth or topography above the reflection point.

observed and calculated arrival times of a seismic wave. In most tomographic studies the Jeffreys–Bullen (JB) model for P-wave velocities²⁹ is used, with earthquake locations and seismic data reported by the ISC without further processing. A disadvantage of the JB model is the smoothness of the velocity–depth profile at depths where upper-mantle discontinuities are known to exist. This difference between JB and the true Earth degrades the linearization and can result in reference-model artefacts at depths of particular interest³⁰. Here we use the iasp91 reference model²⁴, which has rapid increases in P-wave velocity at depths of 410 and 660 km.

There are also good reasons to redetermine the hypocentral parameters published by the ISC. The hypocentres used by the ISC to determine travel-time residuals are computed from P-wave arrival times only. The presence of lateral heterogeneity near the source and the peculiarities of station distribution about most subduction zones can combine to produce a depth bias³¹. Indeed, the large negative mean of reported ISC pP delay times (Fig. 2a and ref. 25) suggests that the focal depths of many subduction zone earthquakes have been overestimated and the associated pP phases potentially misidentified by the ISC (ref. 25, EH91).

For delay-time tomography this problem with ISC hypocentres has two important effects. First, the use of negative reported pP delay times can result in tomographic images that are improperly biased to higher velocities beneath pP reflection points. Second, errors in earthquake hypocentres tend to absorb structurally related signals along P-wave paths near the source, which are not fully recoverable on tomographic inversion.

Before inversion, a combination of ISC P- and pP-wave arrival times and the iasp91 velocity model were used to redetermine earthquake hypocentres in the northwest Pacific region (EH91). New travel-time residuals for reidentified P and pP phases (hereafter referred to as NWP data) were computed from the relocated hypocentres. Figs 2 and 3 demonstrate the effect of the use of iasp91 and hypocentre relocation on the distribution of P and pP residuals. The relocation with the additional information from the pP phase would also lead to reduced variance for pP data even if the JB model were retained as the background Earth model. But for reasons mentioned earlier and because of the large amounts of computer time involved in the relocation procedure, we have relocated directly using the iasp91 model. The improvement is clear from the histograms in Fig. 2.

The variance reduction for P and pP data (Figs 2 and 3) is significant with respect to variance reductions typically obtained

from tomographic inversion^{26,32–34}. This demonstrates that in linearized tomographic inversions of uncorrected ISC data there is a danger of incorrectly mapping inadequacies in the reference model and errors in the earthquake hypocentres into the three-dimensional structure being determined.

Subduction-zone morphology and mantle structure

Jordan and coworkers^{10–13} advocated the penetration of subducted slab to depths of at least 1,200 km below northwest Pacific island arcs. From tomographic studies with P-wave data, some investigators have supported this interpretation^{16,17,34} whereas, others suggested that slab penetration occurs only locally^{19,33,35}. To better understand this disagreement we compare below inversions of reported ISC P data with inversions of our NWP P and pP data (Figs 4 and 5). In our discussion we focus on the structure near the transition from upper to lower mantle.

We inverted delay times for aspherical variations in seismic velocity, for hypocentre relocation owing to near-source heterogeneity, and for station corrections. For the inversion we employed a damped least-squares method based on the LSQR algorithm^{36–39}. We will present the complete results of our inversion elsewhere (R.v.d.H., R.E. & W.S., manuscript in preparation, hereafter referred to as HES91).

The inversion results obtained from reported ISC P data using a JB reference (Figs 4a, b and 5a, b) are characterized by a number of important general features. Higher-than-average P-wave velocities are imaged in upper-mantle regions where subduction of the Pacific plate is known to occur. The images suggest continuity of the high-velocity zones across the 670-km discontinuity, although the amplitude of the lower-mantle velocity variations relative to JB (~2%) are smaller than in the upper mantle (~3%). In the upper mantle, the locations of earthquakes correlate reasonably well with the regions of high P-wave velocity, although the dip of the Wadati–Benioff zone deviates locally from the dip of the imaged high-velocity zone. In the mantle above the seismic zone, amplitudes of aspherical variations of P-wave velocity are low. Locally, an increase in P-wave velocity is imaged near a depth of 670 km.

For upper-mantle regions in the vicinity of the Wadati–Benioff zones, results of tomographic inversions of NWP P and pP data using the iasp91 reference model (Figs 4c, d and 5c, d) do not differ significantly from the ISC images. In detail, however, there are differences in the amplitudes of the velocity variations, and in the correlation between seismicity and zones of high P-wave velocities. The amplitudes of the velocity perturbations

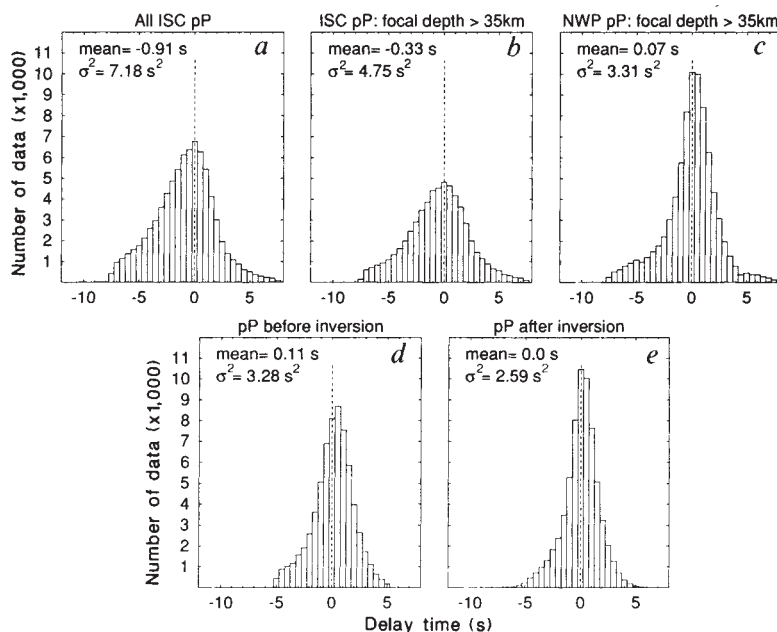


FIG. 2 Frequency distribution of pP data at different stages of data processing. a, pP data reported by the ISC between 1964 and 1987. These data are used only to identify pP data initially, but are not used in our inversions. b, The distribution of ISC pP data from earthquakes with a focal depth larger than 35 km. The maximum value of the travel-time residual is 7.5 s. c, The distribution of NWP pP delay times, which are computed (EH91) from hypocentres that are relocated in the iasp91 model for P-wave velocities²⁴. The maximum value of the travel-time residual is 7.5 s. With respect to the distribution in b, a variance reduction of pP data of 30% was obtained. d, The distribution of NWP pP data after the construction of 66,042 composite rays³⁶ and the deletion of pP data with residuals larger than 5 s. This is the distribution of the pP data before inversion. e, The distribution of pP data after tomographic inversion. On inversion for three-dimensional mantle structure, a further variance reduction of 21% was obtained after 35 iterations of the damped LSQR inversion algorithm^{36,37}.

near the deepest earthquakes of the Wadati-Benioff zones are substantially higher in the NWP than in the ISC images.

Differences between the NWP and ISC images are substantial in mantle regions where aspherical Earth structure is not indicated by seismicity. The inversion of ISC P-wave data reveals a zone of moderately high P-wave velocity continuous to lower-mantle depths in all cross-sections. In contrast, our new results indicate that fast slab-like regions in the lower mantle are found only below the deepest earthquakes of the Kuril and Mariana subduction zones (Figs 5c, d and 6a). The new images suggest, however, a maximum depth of about 850 km for the high-velocity zone below the northern Kuril arc (Figs 5c and 6a), whereas no maximum depth was indicated from the inversion of ISC P-wave data (Fig. 5a). Figure 6a suggests the broadening of the Kuril slab across the transition from upper to lower mantle. The vertical mantle sections across the northern part of the Philippine Sea Plate (Figs 4c and 6c), the Sea of Japan (Fig. 4d) and the southern Kuril arc (Fig. 6b) show images of fast slab-like zones deflected at depths between 500 and 670 km. Figure 6b and c illustrates upper to lower mantle. The relationship between the images of high P-wave velocities and unusual locations of earthquakes several hundreds of kilometres off the inclined Wadati-Benioff zones^{40,41}. Okino *et al.*⁴⁰ explained observed P-wave arrival times from an earthquake off the Izu Bonin seismic zone (Fig. 6c) by a model of a deflected slab with a P-wave propagation 3.5% faster than in the JB model. This is in good agreement with the results from our tomographic imaging (Figs 4c and 6c).

Vertical velocity contrasts at a depth of 670 km are absent in the NWP images. This was expected, as we corrected for the presence of the 670-km discontinuity by using iasp91 rather than

JB as the reference model (compare Figs 5b and d). The section across the southern Philippine Sea (Fig. 5c), however, shows a rapid increase in P-wave velocity between the 410- and 670-km discontinuities. Below the northern part of the Philippine Sea and the Sea of Japan, a comparable increase in P-wave velocity at a depth of about 500 km coincides with the upper boundary of the deflected continuations of the slab-like structures below the northern Izu Bonin, Japan and southern Kuril arcs (Figs 4c, d and 6b). These observations suggest that a seismic discontinuity at a depth near 520 km (ref. 42) is associated at least locally with subducted lithosphere deflected at the base of the upper mantle⁴³.

Reliability of the tomographic images

The reliability of results of large-scale inversions is assessed by calculating the response to a synthetic input model^{8,33,36,44}. But these resolution tests do not reflect inadequacies of the reference model, systematic errors in the data, inconsistencies between travel times and earthquake locations and ray paths, and biases in earthquake location. As a consequence, the effect on resolution of the use of better data, earthquake locations and reference model cannot be demonstrated by resolution tests. But in regions where the sampling by P-wave ray paths only is inadequate, significant improvement in computed resolution is obtained with the additional ray-path coverage and focal-depth resolution provided by the pP-wave data (HES91). This applies in particular to mantle structure beneath intraplate regions and to deflected aseismic parts of the subducted slab. In studies in which only P-wave ray paths are used, these mantle regions are often overlooked³⁴ or characterized by a lack of horizontal resolution^{19,26,33}.

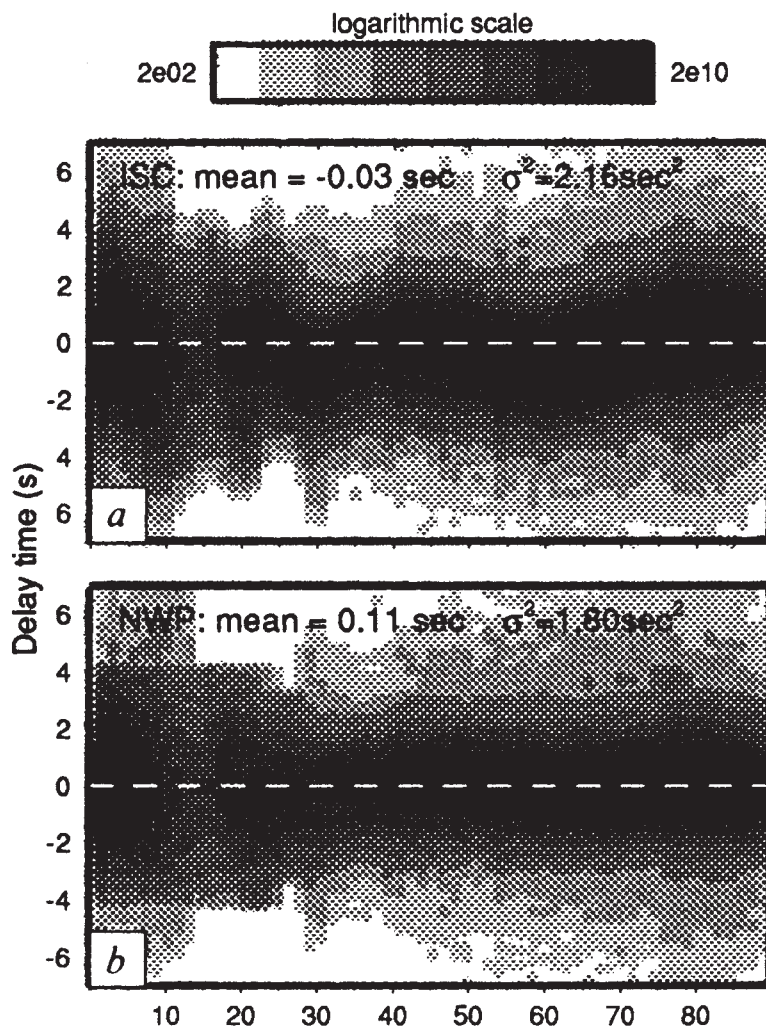


FIG. 3 The distribution of P-wave travel-time residuals versus epicentral distance. *a*, P delay times reported by the ISC (1964–1987) and the National Earthquake Information Center (NEIC, 1987–1989). These data are computed from ISC/NEIC hypocentres and the Jeffreys–Bullen travel-time tables²⁹. Displaying the data in this way shows that waves travelling to stations between 50° and 70° arrive systematically earlier than expected from the JB model. In the distance range of 15° and 25° the concentration of P-wave data suggests the presence of triplication branches owing to upper-mantle discontinuities unaccounted for by the JB model^{30,52}. *b*, Northwest Pacific (NWP) P delay times computed from improved earthquake hypocentres and the iasp91 velocity model²⁴. The NWP P data are evenly distributed around the axis of zero delay time: the branches between 15° and 25° are reduced because by using iasp91 we corrected for upper-mantle discontinuities, and the slightly different lower-mantle structure of iasp91 removed the systematic deviations between 25° and 90°. We remark that the mean value of the ISC P distribution is slightly closer to zero than that of the NWP P data because the ISC data are optimal with respect to ISC hypocentre locations, whereas the NWP P data are computed from hypocentres located with both P and pP data. The variance of NWP P data is 17% less than the variance of the uncorrected ISC P data. On subsequent tomographic inversion, a further variance reduction of 41% was obtained after 35 iterations (HES91).

The result of resolution tests demonstrates that our approach does provide good resolution of lower-mantle structure (HES91). Here we show only the result of a simple test. For the earthquake and station distributions shown in Fig. 1, we computed synthetic P- and pP-wave travel-time residuals through a target anomaly (8% faster than iasp91) at a depth of 975 km. We added gaussian noise to the synthetic data using typical standard errors for P and pP data (1 s and 1.5 s respectively). The (20%) contour lines of the inversion response to this target anomaly are distorted only slightly, which indicates that there is no significant 'smearing' in the dip direction of the slab.

Outside the area under study, we assume the reference Earth structure (here iasp91) and there is a danger that the images are affected by the mapping into the solution of aspherical structure outside the mantle volume under study^{21,22}. The contribution to this mapping is largest from mantle regions in which the Fresnel zone of the P and pP waves decreases in size, that is, in the direct vicinity of the stations outside the study region. We therefore computed station corrections, on tomographic inver-

sion, for all stations outside the study area³⁶. Where available, we used independently determined station corrections⁴⁵ as starting values for our corrections. Studies of the effects of wavefront healing (T. J. Moser and G. Nolet, manuscript in preparation) shows that lower-mantle heterogeneity has a small effect on delay-time variance, whereas systematic bias is mostly absorbed in the station correction.

From the resolution test we concluded that 'smearing' in the dip direction of the Kuril slab is not significant. Together with the observation (Fig. 5c) that low P-wave velocities are imaged below the high-velocity slab (in contrast to Fig. 5a), this leads us to conclude that the high-velocity structure below the Kuril arc does not result from the mapping of distant structure. Also, from analysis of the solution itself we can be confident that no important artefacts from heterogeneities elsewhere have been introduced into the model. Figure 6a and b shows images across the Kuril arc (for locations see Fig. 1). If aspherical structure outside the mantle volume under study contributes to the imaged high P-wave velocity in the lower mantle below the

FIG. 4 Vertical mantle sections across the Izu Bonin and Central Japan island arcs (locations are given in Fig. 1b). The seismicity of the subduction zones is shown above the tomographic images. The horizontal line depicts the 670-km discontinuity. *a*, Izu Boni cross-section by inversion of ISC P data. Note the bent high-velocity zone across the upper-lower mantle boundary below the Izu Bonin arc. This structure was also reported in ref. 34. *b*, Central Japan cross-section by inversion of ISC P data. The 'smearing' of anomalies from the bottom of the upper mantle to the surface reveals lack of resolution along P-wave ray paths to stations in Asia²⁶. *c*, Izu Bonin cross-section by inversion of NWP P and pP data. *d*, Central Japan cross-section by inversion of NWP P and pP data. The tomographic images shown here and in Fig. 5 are obtained after 30 iterations of the damped LSQR inversion algorithm^{36,37}. This algorithm has intrinsic damping properties that minimize artefacts in the images of mantle structure in regions with low resolution. We inverted travel-time residuals for variations in P-wave velocity relative to the one-dimensional reference velocity models shown at the right-hand side of the figures. The Earth's volume under study was parameterized by 47,850 blocks with horizontal dimensions of $1^\circ \times 1^\circ$ and thickness varying from 35 km for the uppermost layer to 200 km at a depth of 1,500 km (HES91).

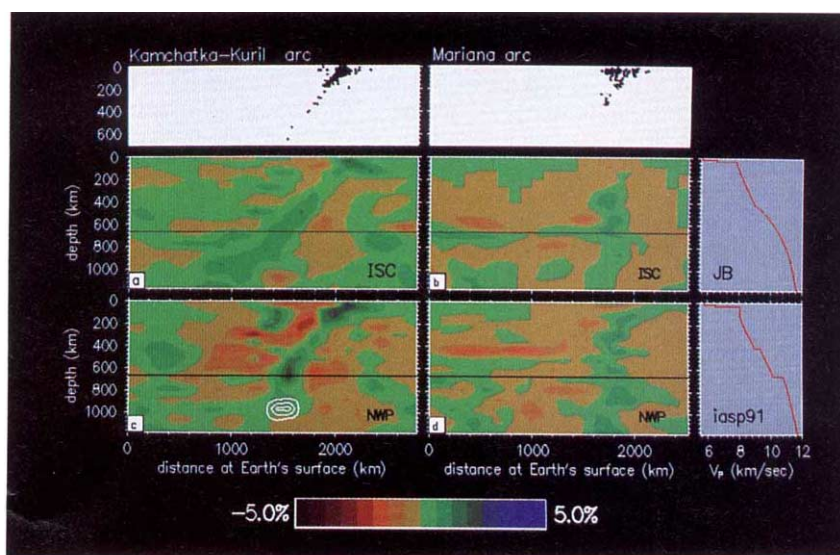
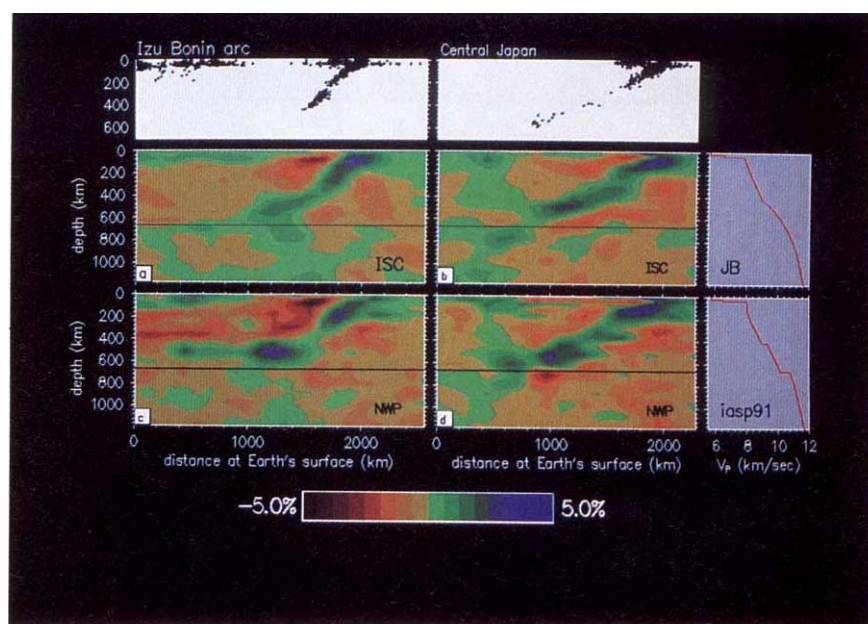


FIG. 5 Mantle sections across the Kuril and Mariana island arcs (locations given in Fig. 1). *a*, Kuril cross-section by inversion of ISC P data. *b*, Mariana cross-section by inversion of ISC P data. Note the absence of high P-wave velocities near the shallow part of the Wadati-Benioff zone. *c*, Kuril cross-section by inversion of NWP P and pP data. *d*, Mariana cross-section by inversion of NWP P and pP data. In *c* we also show the result of a simple resolution test (see text for discussion).

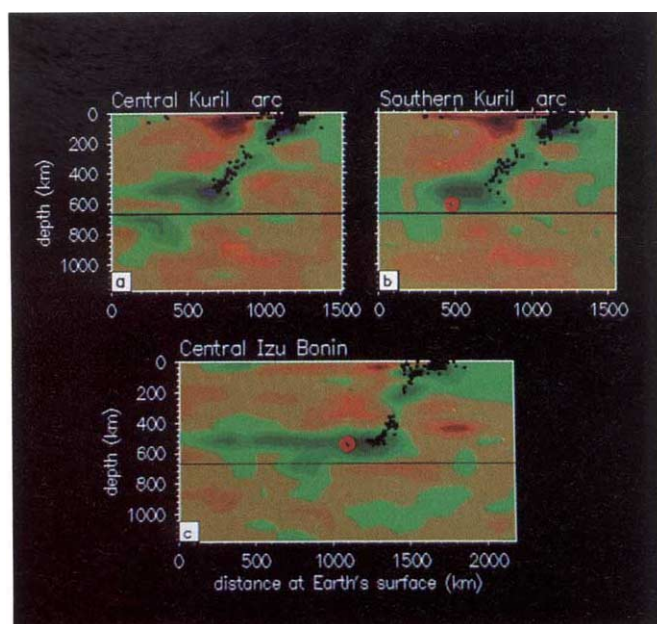


FIG. 6 Mantle cross-sections showing images of slab structure by the inversion of NWP P and pP data (locations given in Fig. 1). a, Central Kuril. b, Southern Kuril, across the island of Sakhalin. The red dot depicts the location of the earthquake that occurred on 12 May, 1990 (ref. 40). c, Central Izu Bonin. The red dot depicts the July 1982 earthquake⁴¹ ($m_b = 6$).

Wadati-Benioff zone, one would expect the effect to be similar for the two images, as the sampling by seismic ray paths is about the same. This is clearly not the case.

Our images of a sub-horizontal high-velocity structure below the northern Philippine Sea Plate and the Sea of Japan are in agreement with the results of Fukao *et al.*³⁵, who solved simultaneously for updates of the one-dimensional reference model and for aspherical variations in seismic velocity. This indicates that our results are not artefacts of the iasp91 model. Slab deflection at the 670-km discontinuity was also suggested by Zhou and Clayton³³, but they used only a few teleseismic rays and would thus have missed steeply dipping structures in the lower mantle.

Mantle convection and subduction history

The inference that subducted slabs are deflected in the transition zone (410–670 km) below the southern Kuril, Japan and Izu Bonin arc (Figs 4c, d and 6b, c), and sink into the lower mantle below the Kuril–Kamchatka and Mariana arcs (Figs 5c, d and 6a) is in good agreement with other images obtained recently³⁵. These results suggest that the boundary between upper and lower mantle acts as a strong barrier for mantle flow. The observation that (at least locally) slabs sink into the lower mantle is, however, inconsistent with a model of strictly independent convection in the upper and lower mantle.

Seismological observations and numerical simulations of mantle flow suggest that the lower mantle is at least 10–30 times more viscous than the upper mantle⁴⁶. Several studies of mantle convection (for a review see ref. 1) have considered variable degrees of isochemical and compositional density and viscosity contrasts between the upper and lower mantle and have shown that slab deflection and slab penetration might coexist^{47–49}. A recent numerical modelling study⁴⁹ of mantle flow included compressibility and the endothermic phase change from spinel to perovskite and magnesiowustite which is associated with the seismic discontinuity at a depth of 670 km. For a shallow negative Clapeyron slope of $-2 \text{ MPa } ^\circ\text{C}^{-1}$ they found intermittent mixing between the upper and lower mantle with flow structures resembling our images of subducted slab below the northwest Pacific.

Our images are consistent with a convection model that allows local mass transport across the boundary between upper and lower mantle. One should be cautious of such interpretations, however: tomographic images depict variations in seismic

velocity, and from the images alone one can not unambiguously discriminate between actual mass transport across the upper–lower mantle boundary or velocity perturbations in the lower mantle resulting from thermal coupling in layered convection⁵⁰, possibly combined with a locally depressed 670 km discontinuity.

Present convection models fail to explain why subducted lithosphere deflects at the base of the upper mantle below particular island arcs, but sinks into the lower mantle below others. Results of fluid-dynamical modelling⁵¹ suggest that subduction history plays an important role in controlling the morphology of the subducted slab near the upper–lower mantle boundary. Slab deflection at a layer with high viscosity can occur in the case of the oceanward migration of the trench owing to the roll-back of the subduction zone, whereas penetration of subducted lithosphere into a high-viscosity lower mantle is possible when the location of the trench is stable^{47,51}. This provides a working model for a tentative explanation of imaged differences in slab morphology below, for instance, the Izu Bonin and Mariana island arcs. Reconstructions of the tectonic history of the Philippine Sea plate⁵³ indicate that, after initial onset ~ 45 Myr ago, the evolution of the subduction zones below the northern and southern parts of the Philippine Sea plate have been different. In the preferred kinematic model of Seno and Maruyama⁵³, the location of the Mariana trench has been stable, possibly triggering slab penetration, but the Izu Bonin trench migrated northwest, over hundreds of kilometres, to its present location southeast of Japan. This would be consistent with the image of a deflected slab below the Izu Bonin arc (Fig. 4c).

A detailed interpretation is beyond the scope of this paper. Nevertheless, it is clear that the integration of results of seismic imaging with information from tectonic reconstructions and fluid-dynamical modelling will ultimately lead to new understanding of the relationship between plate tectonics and mantle convection. $\square$

Received 31 May; accepted 8 August 1991.

1. Olson, P., Silver, P. G. & Carlson, R. W. *Nature* **344**, 209–215 (1990).
2. Jordan, T. H. & Lynn, W. S. *J. geophys. Res.* **79**, 2679–2685 (1974).
3. Jordan, T. H. *Nature* **257**, 745–750 (1975).
4. Lay, T. *Geophys. J. R. astr. Soc.* **72**, 503–516 (1983).
5. Grand, S. P. *J. geophys. Res.* **92**, 14065–14090 (1987).
6. Vidale, J. E. & Garcia-Gonzalez, D. *Geophys. Res. Lett.* **15**, 369–372 (1988).
7. van der Hilst, R. D. thesis, Univ. of Utrecht (1990).
8. Inoue, H., Fukao, Y., Tanabe, K. & Ogata, Y. *Phys. Earth planet. Inter.* **59**, 294–328 (1990).

9. Brooks, D. A. *et al.* *Tectonophysics* **102**, 1–16 (1984).
10. Jordan, T. H. *J. geophys. Res.* **43**, 473–496 (1977).
11. Creager, K. C. & Jordan, T. H. *J. geophys. Res.* **89**, 3031–3049 (1984).
12. Creager, K. C. & Jordan, T. H. *J. geophys. Res.* **91**, 3573–3589 (1986).
13. Fischer, K. M., Jordan, T. H. & Creager, K. C. *J. geophys. Res.* **93**, 4773–4783 (1988).
14. Silver, P. G. & Chan, W. W. *J. geophys. Res.* **91**, 13787–13802 (1986).
15. Cormier, V. F. *J. geophys. Res.* **94**, 3006–3024 (1989).
16. Suetsugu, D. *J. Phys. Earth* **37**, 265–295 (1989).
17. Takei, Y. & Suetsugu, D. *J. Phys. Earth* **37**, 225–231 (1989).
18. Dziewonski, A. M. *J. geophys. Res.* **89**, 5929–5952 (1984).
19. Zhou, H. *Geophys. Res. Lett.* **15**, 1425–1428 (1988).
20. Lay, T. & Young, C. J. *Geophys. Res. Lett.* **16**, 605–608 (1989).
21. Schwartz, S. Y., Lay, T. & Grand, S. *Geophys. Res. Lett.* **18**, 1265–1268 (1991).
22. Zhou, H., Anderson, D. L. & Clayton, R. W. *J. geophys. Res.* **95**, 6799–6827 (1990).
23. Aki, K., Christofferson, A. & Husebye, E. S. *J. geophys. Res.* **82**, 277–296 (1977).
24. Kennett, B. & Engdahl, E. R. *Geophys. J.* **105**, 429–465 (1991).
25. van der Hilst, R. D. & Engdahl, E. R. *Geophys. J.* **106**, 169–188 (1991).
26. Spakman, W., Stein, S., van der Hilst, R. D. & Wortel, R. *Geophys. Res. Lett.* **16**, 1097–1101 (1989).
27. Engdahl, E. R. & Billington, S. *Bull. seismol. Soc. Am.* **76**, 77–93 (1986).
28. Engdahl, E. R. & Gubbins, D. *J. geophys. Res.* **92**, 13855–13862 (1987).
29. Jeffreys, H. & Bullen, K. E. *British Association for the Advancement of Science* (London, 1967).
30. van der Hilst, R. D. & Spakman, W. *Geophys. Res. Lett.* **16**, 1093–1096 (1989).
31. Dziewonski, A. & Anderson, D. L. *J. geophys. Res.* **88**, 3295–3314 (1983).
32. Hirahara, K. & Mikumo, T. *Phys. Earth planet. Inter.* **21**, 109–119 (1980).
33. Zhou, H. & Clayton, R. W. *J. geophys. Res.* **95**, 6829–6851 (1990).
34. Kamiya, S., Miyatake, T. & Hirahara, K. *Geophys. Res. Lett.* **15**, 828–831 (1988).
35. Fukao, Y., Nenbai, M., Obayashi, M. & Inoue, H. *J. geophys. Res.* (submitted).
36. Spakman, W. & Nolet, G. in *Mathematical Geophysics* (eds Vlaar, N. J. *et al.*) 155–188 (Reidel, Dordrecht, 1988).
37. Paige, C. C. & Saunders, M. A. *ACM Trans. Math. Soft.* **8**, 3–71; 195–209 (1982).
38. Nolet, G. *J. comput. Phys.* **61**, 463–482 (1985).
39. van der Sluis, A. & van der Vorst, H. A. in *Seismic Tomography* (ed. Nolet, G.) 49–84 (Reidel, Dordrecht, 1987).
40. Okino, K., Ando, M., Kaneshima, S. & Hirahara, K. *Geophys. Res. Lett.* **16**, 1059–1063 (1989).
41. Ekström, G., Dziewonski, A. & Ibañez, J. *Eos* **71**, 1462 (1990).
42. Shearer, P. *Nature* **344**, 121–126 (1990).
43. Anderson, D. L. *Eos* **71**, 1473 (1990).
44. Humphreys, E. & Clayton, R. W. *J. geophys. Res.* **93**, 1073–1085 (1988).
45. Toy, K. thesis, Univ. of California, San Diego (1990).
46. Gurnis, M. & Hager, B. *Nature* **335**, 317–321 (1988).
47. Ringwood, A. E. & Irifune, T. *Nature* **331**, 131–136 (1988).
48. Christensen, U. & Yuen, D. A. *J. geophys. Res.* **89**, 4389–4402 (1984).
49. Machetel, P. & Weber, P. *Nature* **350**, 55–57 (1991).
50. Nataf, H. C., Moreno, S. & Cardin, P. *J. Phys. France* **49**, 1707–1714 (1989).
51. Kincaid, C. & Olson, P. *J. geophys. Res.* **92**, 13832–13840 (1987).
52. Zielhuis, A., Spakman, W. & Nolet, G. in *Digital Seismology and Modelling of the Lithosphere* (eds Cassinis, R., Panza, G. & Nolet, G.) 333–340 (Plenum, London, 1988).

ACKNOWLEDGEMENTS. We thank Brian Kennett for making the iasp models for P and S wave velocity available to us during their development. A review by Thorne Lay greatly improved the manuscript. This research was financially supported by NATO. Tomographic inversions were performed on the C205 supercomputer of the SARA (Amsterdam, The Netherlands) with financial support from the Netherlands Organization for the Advancement of Pure Research.

Assignment of G-protein subtypes to specific receptors inducing inhibition of calcium currents

C. Kleuss*, J. Hescheler†, C. Ewel*, W. Rosenthal†, G. Schultz† & B. Wittig*

*Institut für Molekularbiologie und Biochemie, and †Institut für Pharmakologie, Freie Universität Berlin, Arnimallee 22, D-W1000 Berlin 33, Germany

The inhibition of voltage-dependent Ca^{2+} channels in secretory cells by plasma membrane receptors is mediated by pertussis toxin-sensitive G proteins. Multiple forms of G proteins have been described, differing principally in their α subunits, but it has not been possible to establish which G-protein subtype mediates inhibition by a specific receptor. By intranuclear injection of antisense oligonucleotides into rat pituitary GH_3 cells, the essential role of the G_o -type G proteins in Ca^{2+} -channel inhibition is established: the subtypes G_{o1} and G_{o2} mediate inhibition through the muscarinic and somatostatin receptors, respectively.

REGULATORY GTP-binding proteins (G proteins) constitute a family of proteins involved in signal transduction across the plasma membrane^{1,2}. They consist of three different subunits, α , β and γ . G proteins are classified according to their α subunits, some of which are substrates for ADP-ribosylating bacterial exotoxins, such as cholera and pertussis toxin. Sixteen G protein α subunits encoded by different genes are known, and this diversity of α subunits is increased by alternative splicing². In combination with a particular α subunit, at least three types of β and four types of γ subunits contribute to an even greater variety of heterotrimeric G proteins.

G proteins transduce signals from membrane receptors to effectors like enzymes (for example, adenylyl cyclase and phospholipase C) and ion channels (such as the voltage-sensitive calcium or potassium channels). But unequivocal assignment of one G protein to a single receptor/effector system has been achieved only for G_s , the stimulatory G protein of adenylyl cyclase, and for the transducins, the retinal G proteins, which

mediate between activated photoreceptors and a cyclic GMP phosphodiesterase. As the α subunits are very similar, it is difficult to distinguish between the G-protein subtypes: recombinant G-protein subunits expressed in bacteria are not particularly active, possibly because of inadequate post-translational modification^{3,4}, and antibodies raised against peptides corresponding to specific regions of G-protein subtypes have little affinity for native G proteins.

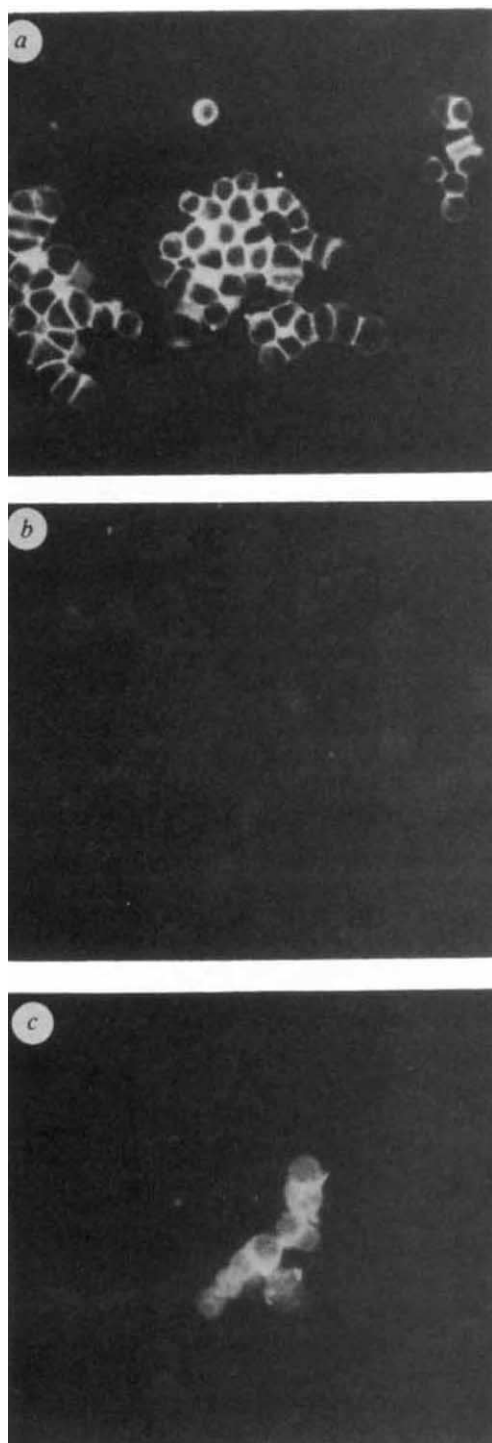
Our study concerns the G_o family, which is limited to neuronal, neuroendocrine and endocrine cells^{5–7}, and makes use of the fact that the distinction between the α subunits of G_o subtypes is greater at nucleic acid level than in the translated protein. We have identified selective base sequences in each α subunit and designed oligodeoxynucleotides in an antisense orientation to the respective messenger RNAs for specific 'knock-out' experiments. We microinjected these oligonucleotides directly into nuclei, the presumed site of action, in controlled amounts without significantly disturbing cell-surface functions, which might happen should they be added in high concentrations to the culture medium.

Using this approach in rat pituitary GH_3 cells, we assign a function to G_o in the receptor-induced inhibition of voltage-sensitive Ca^{2+} channels and show that the subtype G_{o1} specifically mediates the action of carbachol via a muscarinic receptor and that G_{o2} transduces the signal from the somatostatin receptor. As the cytosolic Ca^{2+} concentration is regulated by Ca^{2+} -channel activity, the receptor-induced inhibition of Ca^{2+} channels by somatostatin or carbachol could be crucial for the control of secretion.

Effectiveness of microinjected antisense DNA

To check whether our procedure does suppress G-protein α subunits, we monitored cells by immunofluorescence microscopy following microinjection of antisense DNA. GH_3 cells express α subunits of $\text{G}_i(\alpha_i)$ and two forms each of α subunits of G_i and G_o (refs 8, 9). G-protein α subunits were

visualized by an affinity-purified α_{common} -peptide antibody¹⁰ which recognizes α_s , α_i , α_o , α_z and the α subunit of transducin^{8,11,12}. In non-injected cells, α subunits are associated mainly with the plasma membrane, as indicated by the ring-shaped pattern of fluorescence obtained with the α_{common} -peptide antibody (Fig. 1). We found that α subunits were not detectable one day after injection of the antisense oligonucleotide $\alpha\text{-com}$, which corresponds to the α_{common} -peptide, indicating that expression of G-protein α subunits was reduced or even abolished. On the third day after injection, expression of α subunits had recovered. Similar results were obtained with cells injected with an α_o -specific oligonucleotide after staining with an α_o -specific antibody (data not shown).



G proteins in Ca^{2+} current inhibition

To show that G-protein α subunits are involved in the agonist-induced inhibition of Ca^{2+} currents in rat pituitary GH₃ cells, we made electrophysiological observations after injection of an antisense oligonucleotide common to all α subunits.

Ca^{2+} currents in GH₃ cells are probably due to activation of L-type channels¹² and are inhibited to 70–75% of control currents by carbachol and somatostatin (Fig. 2a) in a non-additive manner. Inhibition is abolished by treating cells with pertussis toxin. As shown previously^{8,12} these Ca^{2+} currents and their inhibition are not affected by intracellularly applied cyclic AMP. It follows that G_i -mediated inhibition of adenyl cyclase is not involved.

After injection of the oligonucleotide $\alpha\text{-com}$ (Fig. 2b), inhibition of Ca^{2+} currents by somatostatin is reduced in a time-dependent way: after 48 h, the peptide hormone is without effect, but after about 72 h the response of Ca^{2+} currents to somatostatin is restored. A similar time course of somatostatin-induced Ca^{2+} current inhibition is observed when cells are injected with the antisense oligonucleotide GS-o, which is complementary to α_o mRNAs (Fig. 2b). As with somatostatin, the carbachol-induced inhibition of Ca^{2+} currents is transiently abolished by each oligonucleotide (results not shown). These results show that the expression of G-protein α subunits, probably α_o , correlates with the ability of receptor agonists to inhibit Ca^{2+} currents.

The involvement of the β subunit in G-protein-mediated Ca^{2+} -current inhibition was studied by microinjecting the oligonucleotide anti- β , which is complementary to the 5' coding regions of β_1 and β_2 mRNAs. It contains 6 nucleotides out of 29 that do not match β_3 mRNA, the longest continuous match being 13 bases. Thus, the anti- β oligonucleotide will probably form stable hybrids with β_3 mRNA (refs 13, 14). Injection of about 10,000 full-length molecules of anti- β oligonucleotide had no major influence on the receptor-induced Ca^{2+} -channel inhibition (Fig. 3). This is in contrast with the same number being fully effective in the case of antisense oligonucleotides against α subunits. If $\sim 50,000$ full-length anti- β oligonucleotide molecules are injected, inhibition of Ca^{2+} currents by somatostatin (Fig. 3) and carbachol (not shown) is abolished. The

FIG. 1 Immunocytochemical detection of G-protein α subunits in GH₃ cells injected with antisense oligonucleotides. a, Non-injected cells; b and c, cells intranuclearly injected with oligonucleotide $\alpha\text{-com}$ and incubated in culture medium for 24 h (b) or 72 h (c). Sequence of injected oligonucleotide $\alpha\text{-com}$: TCATYTGCTTCAATGGTRCTYTTYCCRGATTC, corresponding to nucleotides 133–160 of the identical strand of the α_{o2} gene sequence³⁹; abbreviations for wobbled positions: R (G or A), Y (T or C), M (A or C), K (G or T), S (G or C), W (A or T)⁴⁰. Oligonucleotide $\alpha\text{-com}$ can hybridize with mRNAs of all known variants of α_i , α_o , α_s and with the α subunits of transducin.

METHODS. GH₃ cells were cultured for 5 to 9 days on glass slides imprinted with numbered squares for convenient localization of the cells injected^{8,12,41}. Sequences of selective oligonucleotides were detected with the help of sequence comparison and multiple alignment of the MacMolly program (Soft Gene GmbH, Berlin). Oligonucleotides were synthesized on a DNA synthesizer (Milligen Model 8600) or purchased (TIB Molbiol, Berlin). After cleavage of the protecting groups, DNA was extracted with phenol and chloroform and precipitated with ethanol. Intranuclear injections of oligonucleotides (5 μM) were performed with an automated injection system (Zeiss). No difference in effectiveness of injected oligonucleotides could be detected if they were dissolved in a physiological nuclear buffer ('new' buffer⁴²) instead of water. About 3,000 cells were injected by means of commercial microcapillaries (Femtotips, Eppendorf) with an outlet diameter of 0.2 μm . Injection time was 0.1 s, the pressure 20–30 hPa. The calculated injected volume was 10–20 fl (ref. 43) containing 10,000 to 20,000 full-length oligonucleotides. Indirect immunofluorescence staining was carried out as described⁴⁴. After fixation cells were incubated with the first antibody overnight at 6–8 °C. The affinity-purified α_{common} -peptide antibody, which recognize the α subunits of α_s , α_i , α_{i2} , α_{i3} , α_{o1} , α_{o2} , α_z and transducin, was diluted 100-fold in phosphate-buffered saline. Incubation with the second antibody, 1:100 diluted goat anti-rabbit IgG-fluorescein isothiocyanate conjugate (Sigma), was for 1 h at room temperature. Scale bar, 20 μm .

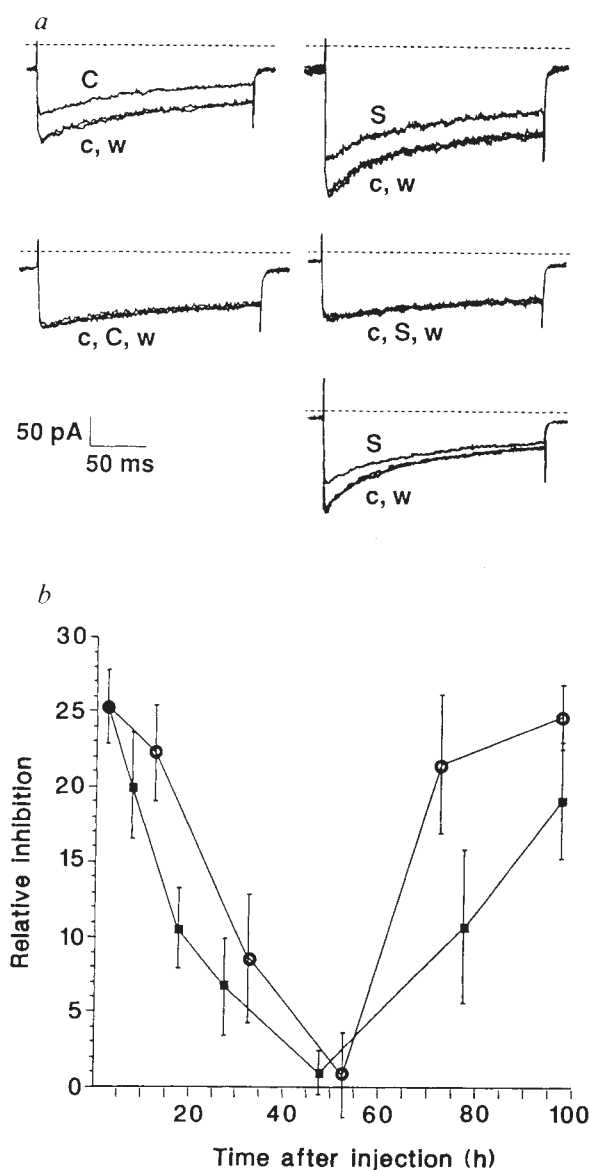
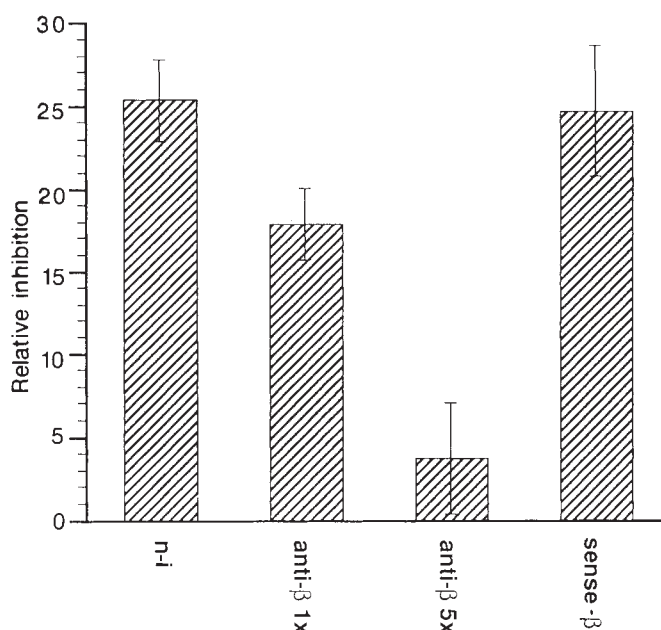


FIG. 3 Ca^{2+} -current inhibition by somatostatin in GH_3 cells injected with antisense oligonucleotides directed against mRNAs encoding β subunits. Whole-cell Ca^{2+} currents were measured 24 h after injection of the oligonucleotide indicated. The inhibition of Ca^{2+} currents by somatostatin ($1 \mu\text{M}$) is shown in per cent of currents observed in the absence of the receptor agonist (mean values are given; $n \geq 8$). n-i, Non-injected cells; anti- β 1 $\times$ and anti- β 5 $\times$, about 10,000 and 50,000 full-length molecules respectively, were injected per nucleus. Cells that had been injected with anti- β 1 $\times$ responded to somatostatin in the same way as non-injected cells ($P < 0.01$, Kolmogoroff-Smirnoff test). Sequences of injected oligonucleotides: anti- β , GTSCCATYTGATTCTYCCACWGGGTC, corresponding to nucleotides 112–140 of the identical strand of the β_1 gene sequence⁴⁶. It can hybridize with the three known sequences of β mRNAs (see text); sense- β , GACCCWGTGGGRAGAATCCARATGMSGAC, corresponding to nucleotides 112–140 of the identical strand of the β_1 gene sequence⁴⁶. It has reverse complementarity to anti- β . Abbreviations for wobbled positions are given in the legend to Fig. 1.

FIG. 2 a, Ca^{2+} -current inhibition by somatostatin and carbachol in GH_3 cells injected with antisense oligonucleotides. Whole-cell Ca^{2+} currents were recorded under voltage-clamp conditions by depolarizing pulses from -80 to 0 mV. c, Control currents before application of receptor agonist; S, currents during superfusion of cells with $1 \mu\text{M}$ somatostatin; C, currents during superfusion of cells with $10 \mu\text{M}$ carbachol; w, currents after removal of receptor agonists. Top: non-injected cell; middle: cell injected with GS-o; bottom: cell injected with sense- α . b, Time course of Ca^{2+} -current inhibition by somatostatin ($1 \mu\text{M}$) in GH_3 cells injected with antisense oligonucleotides. Oligonucleotides α -com (filled squares) and GS-o (open circles) were injected at time zero. Mean values are shown ($n \geq 10$). Sequences of injected oligonucleotides: GS-o, CGCCTTGCCGCTCGAG, corresponding to nucleotides 60–78 of the identical strand of the α_{o2} gene sequence³⁹. GS-o can hybridize with mRNAs of all known variants of α_o ; α -com, see legend to Fig. 1; sense- α , GAATCYGGRAARAGYACCATTTGTGAAGCARATGA, corresponding to nucleotides 133–160 of the identical strand of the α_{o2} gene sequence³⁹. It has reverse complementarity to α -com. For abbreviations denoting wobbled positions, see legend to Fig. 1.

METHODS. Cells were grown on coverslips until they reached a density of 500–1,000 cells per frame ($1 \times 1 \text{ mm}^2$). Oligonucleotides were injected into all cells within one frame. For determination of Ca^{2+} currents, coverslips were transferred to a chamber (0.2 ml) mounted on an inverted microscope. The chamber was continuously perfused with a solution with Ba^{2+} as charge carrier, containing 125 mM NaCl, 10.8 mM BaCl_2 , 1 mM MgCl_2 , 5.4 mM CsCl, 10 mM glucose and 10 mM Na-HEPES (pH 7.4 at 37°C); flow rate was 5 ml min^{-1} . The pipette solution contained 120 mM CsCl, 1 mM MgCl_2 , 3 mM Mg-ATP, 10 mM EGTA, 10 mM Na-HEPES (pH 7.4). Whole-cell membrane currents were measured according to ref. 45 using a List/EPC7 patch-clamp amplifier. The holding potential was -80 mV. Inward currents were recorded during 200-ms long test potentials to 0 mV; stimulation rate was 0.5 Hz . The shape, amplitude and inactivation characteristics of Ca^{2+} currents were not affected by microinjection itself. In some cells a rapid inactivating component was eliminated by those antisense oligonucleotides that suppress the agonist-induced inactivation of the voltage-sensitive Ca^{2+} current.



effects of anti- β oligonucleotides are specific, because the corresponding sense oligonucleotide, sense- β , has no effect. Our results emphasize the requirement for β subunits in the interaction between receptor and G protein.

G_o inhibits Ca²⁺ currents

To determine whether the loss of Ca²⁺-current inhibition is due to the suppression of α_o , we tested the ability of various antisense oligonucleotides to interfere with the receptor-induced inhibition of Ca²⁺ currents. As we have already shown, oligonucleotide GS-o, which is complementary to the translated 5'-terminal regions of α_o mRNAs, largely reduces the somatostatin-induced inhibition of Ca²⁺ currents (Fig. 4). In contrast, the antisense oligonucleotides complementary to translated 5'-terminal regions of α_s and α_i mRNAs, GS-s and GS-i respectively, have no effect. The sense oligonucleotide (sense- α) corresponding to α -com, or an oligonucleotide with an unrelated sequence (nonsense) are likewise without effect. Similar results are obtained if carbachol is used instead of somatostatin (data not shown).

We also injected antisense oligonucleotides complementary to sequences located immediately upstream from the translation start as the diversity of α -subunit mRNAs is more pronounced in the leader sequence than in the coding regions. This diversity permitted the use of long oligonucleotides (>20 nucleotides) which form stronger hybrids with the target sequence and should retain the ability to hybridize even after removal of bases by nuclear exonucleases. As shown in Fig. 4, the effects of antisense oligonucleotides that can hybridize with 5'-untranslated regions of α_s mRNA (5's2) and α_o mRNA (5'o3) are similar to the those of oligonucleotides that hybridize with the translated regions of the corresponding mRNAs.

Our data confirm that G_o is involved in the receptor-induced inhibition of voltage-sensitive Ca²⁺ channels. This hypothesis was based on two observations made in neuronal cells: (1) the ability of purified α_o to restore the receptor-induced inhibition of Ca²⁺ currents in pertussis toxin-treated cells^{6,15-17}, and (2) the attenuation of the inhibitory modulation of Ca²⁺ currents by α_o antibodies^{16,18}.

Differential functions of G_o subtypes

Alternative splicing of a single transcript encoding α_o subunits¹⁹ yields three mRNAs, α_{o1A} , α_{o1B} , and α_{o2} (refs 20, 21). As α_{o1A} and α_{o1B} mRNAs differ only in the 3' noncoding regions, they encode the same protein, namely the α_{o1} subunit. The α_{o2} mRNA encodes the α_{o2} subunit. The two α_o subunits differ only in their C-terminal regions, which are assumed to be the recognition site for activated receptors^{22,23}. After injection of GH₃ cells with

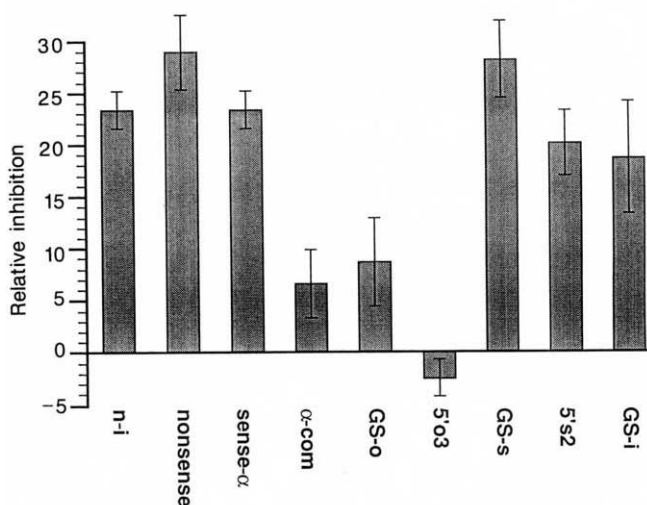
oligonucleotide to1, which has antisense orientation to a 3' coding region of α_{o1} mRNA, the Ca²⁺-current inhibition by carbachol is considerably reduced (Fig. 5), in much the same way as after injection of oligonucleotide GS-o; however inhibition induced by somatostatin is not affected. Oligonucleotide to1 should not hybridize with α_{o2} mRNA, but antisense oligonucleotide to2 does. Oligonucleotide to2 is complementary to the corresponding region of α_{o2} mRNA and reduces the ability of somatostatin but not of carbachol to inhibit Ca²⁺ currents. The data indicate that carbachol and somatostatin inhibit Ca²⁺ currents by different G_o subtypes, that is, by G_{o1} and G_{o2} respectively. To exclude the possibility that antisense oligonucleotides against the 3' coding regions cause a selective suppression of hormone responses irrespective of their sequence specificity for one of the α_o mRNAs, we used another oligonucleotide, to3, which is complementary to a 3' coding region common to all known α_o mRNAs. As shown in Fig. 5a, this antisense oligonucleotide reduces the sensitivity of cells to each agonist.

The different functions of G_{o1} and G_{o2} were confirmed by microinjecting oligonucleotides complementary to the individual 3' nontranslated regions of the various α_o mRNAs (Fig. 5). Antisense oligonucleotides 3'o1A and 3'o1B both abolish the carbachol-induced inhibition of Ca²⁺ currents, whereas oligonucleotide 3'o2, complementary to α_{o2} mRNA, only abolishes the response of cells to somatostatin. These results taken together demonstrate that the α_{o1} subunit of the G-protein trimeric complex mediates the action of the carbachol receptor, whereas the α_{o2} subunit mediates that of the somatostatin receptor. Alternatively spliced products of a single gene thus have distinct functions, which are presumably defined by their C-terminal regions^{22,23}.

Discussion

Our results support the hypothesis⁶ that G_o-type G proteins, but not members of the G_i family, mediate between inhibitory receptors and voltage-sensitive Ca²⁺ channels. Injection of antisense oligonucleotides into nuclei of rat pituitary GH₃ cells largely inhibits the expression of G-protein subunits, and oligonucleotides complementary to mRNAs encoding the two α_o subunits remove the inhibition of voltage-dependent Ca²⁺ currents by carbachol and somatostatin. This modulates the intracellular Ca²⁺ concentration and may thereby alter the secretion behaviour of the cell. Our data from GH₃ cells are consistent with an effect observed in rat pituitary tumour cells, in which expression of α_o was correlated with the inhibition of prolactin secretion by dopamine²⁴. Oligonucleotides complementary to

FIG. 4 Ca²⁺-current inhibition by somatostatin in GH₃ cells injected with antisense oligonucleotides directed against mRNAs encoding different types of G-protein α subunits. Whole-cell Ca²⁺ currents were measured about 24 h after injection of the oligonucleotide. The inhibition of Ca²⁺ currents induced by somatostatin (1 μ M) is shown in per cent of currents in the absence of the receptor agonist (mean values; $n \geq 10$). n-i, Non-injected cells. Cells that had been injected with 5's2 or GS-i responded to somatostatin in the same way as non-injected cells ($P < 0.01$, Kolmogoroff-Smirnoff test). Sequences of injected oligonucleotides: nonsense, GGGGAAGTAGG-TCTTGGTGGTGGG; α -com, see legend to Fig. 1; sense- α and GS-o, see legend to Fig. 2; 5'o3, GGTGGCCCCCTCCCTGCCACAGCCGCGAGACTCG, corresponding to nucleotides (-35)-(-1) of the identical strand of the α_o gene sequence⁴⁷. It is able to hybridize with the described α_o mRNA; GS-s, TTGTGGCCTCROGCTG, corresponding to nucleotides 54-70 of the identical strand of the α_s gene sequence⁴⁷. It can hybridize with all known mRNA sequences coding for G_s-type G proteins; 5's2, GCGGGCGCGGGCGCGGC-CGGGCTGCGGGGCGGCG, corresponding to nucleotides (-36)-(-1) of the identical strand of the α_s gene⁴⁷. It can hybridize with the α_s mRNA described; GS-i, ARGTTSYKGTGATCAT, corresponding to nucleotides 51-67 of the identical strand of the α_i gene sequence⁴⁷. It can hybridize with all known mRNA sequences coding for G_i-type G proteins. Abbreviations for wobbled positions are given in the legend to Fig. 1.



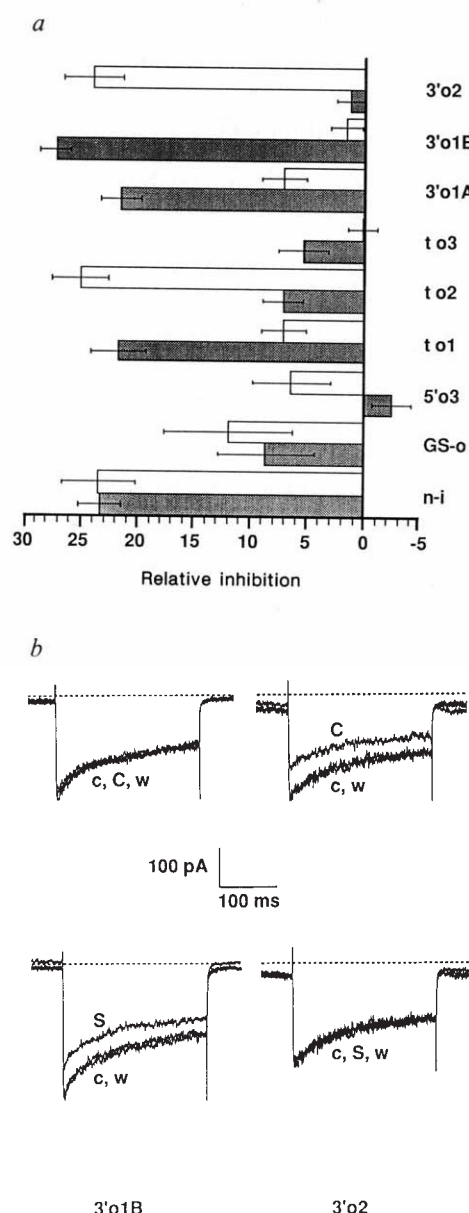


FIG. 5 Ca^{2+} current inhibition by hormones in GH₃ cells injected with antisense oligonucleotides directed against mRNAs encoding α_o . Whole-cell Ca^{2+} currents were measured about 24 h after injection of the oligonucleotide. **a**, Inhibition of Ca^{2+} currents by somatostatin (1 μM ; shaded bars) or carbachol (10 μM ; open bars) as a per cent of the currents observed in the absence of the respective agonist (mean values; $n \geq 10$). n-i, Non-injected cells. **b**, Time-current recordings of the voltage-sensitive Ca^{2+} channel in GH₃ cells in the presence of carbachol (upper) or somatostatin (lower). Left, one cell injected with 3'o1B; right, one cell injected with 3'o2. See legend of Fig. 2 for abbreviations. Sequences of injected oligonucleotides: nonsense, see legend to Fig. 4; GS-o, see legend to Fig. 2; 5'o3, see legend to Fig. 4; to1, AGGCAGCTGCATCTTCATAGGTGTT, corresponding to nucleotides 907–930 of the identical strand of the α_{o1} gene sequence³⁹. It can hybridize with α_{o1} mRNA; to2, GAGCCACAGCTTCTGTGAAGGCACT, corresponding to nucleotides 907–930 of the identical strand of the α_{o2} gene sequence³⁹. It can hybridize with α_{o2} mRNA; to3, TGGGCTGCTGATTTCAGG, corresponding to nucleotides 888–906 of the identical strand of the α_{o1}/α_{o2} gene sequence²¹. It can hybridize with both α_o mRNAs; 3'o1A, TCAACAGCAAAGAGTCCATGAAGCAGT, corresponding to nucleotides 1,093–1,119 of the identical strand of the 'G_oA α clone G_o12' sequence²¹. It can hybridize with the described α_{o1A} mRNA; 3'o1B, CAAGCTGTCTGTGAGTGTGGAAT, corresponding to nucleotides 1,096–1,120 of the identical of the 'G_oA α clone G_o11' sequence²¹. It can hybridize with the described α_{o1B} mRNA; 3'o2, CAGAGCTCTGGGTCCAGGGGAGAAGTG, corresponding to nucleotides 1,093–1,119 of the identical strand of the α_{o2} gene sequence²¹. It can hybridize with α_{o2} mRNA. Abbreviations for wobbled positions are given in the legend to Fig. 1.

α_{o1} mRNA selectively interfered with Ca^{2+} -current inhibition by carbachol, whereas oligonucleotides complementary to α_{o2} mRNA specifically abolished the somatostatin-induced inhibition of Ca^{2+} currents. Intranuclear injection of oligonucleotides complementary to the mRNAs encoding G-protein β subunits also prevented Ca^{2+} -current inhibition. This demonstrates the absolute requirement of β subunits for effective coupling between receptor and G protein.

The successful repression of G protein function implies that the inhibition of synthesis lasts longer than the lifetime of subunits already expressed. The half-lives of α_o subunits range from 21 h to 35 h under steady-state conditions²⁵. In our experiments, immunofluorescence of G-protein α subunits was barely detectable 24 h after injection of oligonucleotides (Fig. 1). Function was impaired 19 h after injection of oligonucleotides (Fig. 2). This discrepancy can be explained by the disturbance of gene regulation or of steady-state levels by the injection of oligonucleotides, leading to an altered lifetime of G-protein subunits.

The selective repression of individual genes by nucleic acids in antisense orientation, that is, complementary to their mRNAs, has often been demonstrated²⁶ (for review, see refs 27, 28). If antisense DNA oligonucleotides are delivered to recipient cells, they hybridize to complementary regions in mRNA and render it susceptible to cellular RNase activity^{29,30}.

Conventional techniques for delivering antisense nucleic acids into mammalian cells are not suitable for suppressing G protein α subunits: only a few short sequence regions allow for the selective hybridization of antisense oligonucleotides. Such short regions cannot be transcribed either transiently or permanently into antisense RNA from recombinant eukaryotic expression vectors. Alternatively, short DNA oligonucleotides can be delivered to the cytoplasm of mammalian cells by addition to the culture medium at high concentration. But the efficiency of uptake and the amount of oligonucleotide surviving passage through degradative organelles is influenced by the culture medium, the cell type, the stage of the cell cycle, the length of oligonucleotide and probably its sequence³¹. Moreover, incubating cells at micromolar concentrations of an oligonucleotide might interfere with cell-surface functions. Taking these considerations into account, we delivered our DNA oligonucleotides into nuclei by microinjection, a technique virtually independent of oligonucleotide length and sequence: the oligonucleotides are applied to the presumed site of action, the amount reaching the nucleus is controlled, and the optimal culture environment is unchanged.

Although we do not consider here the mechanism of antisense repression of genes (for review, see ref. 32), our results might help to specify the target level. A translational block due to a hybridization complex between antisense DNA oligonucleotide and mRNA is very unlikely to be involved. Depending on the mRNA region, such a complex could either interfere with translation initiation or lead to premature termination as a result of stalling by ribosomes. Our data show that the repression is independent of assumed hybridization positions. Oligonucleotides against untranslated sequences upstream or downstream from coding sequences are equally effective in repressing α subunits. In particular, antisense oligonucleotides hybridizing in 3' untranslated regions of α_o mRNA (3'o1A, 3'o1B, 3'o2) cannot cause premature termination.

Considering quantitative conditions, we routinely injected $\sim 10^4$ antisense oligonucleotide molecules into each nucleus: the same degree of repression was found with 10^3 molecules (results not shown), but when 10^2 molecules were injected per nucleus, the repression disappeared. Assuming a steady-state amount of 10 molecules of α_o mRNA per GH₃ cell, an initial 100-fold excess of oligonucleotide over mRNA is sufficient to reduce expression. After intranuclear injection, the number of free oligonucleotide molecules per cell will be reduced by cell divisions, degradation by single strand-specific DNA

exonucleases and by the turnover of the mRNA population. The very low number of molecules necessary for repression can be explained if the hybridization structure inhibits the chromatin template function or if antisense oligonucleotides are continuously re-used after degradation of transcripts. The selective repression of splice variants α_{o1} and α_{o2} argues against the former mechanism. Thus continuous cycles of hybrid formation, the cutting of transcripts at the hybrid site by RNase H, the liberation of antisense oligonucleotide, and hybridization with another α_o transcript are the most likely mechanisms¹³. If the concentration of free antisense oligonucleotide is low and hybridization complexes are the predominant structures, oligonucleotides are protected from degradation.

Antisense oligonucleotides can be used to study the role of any G-protein subunit: for instance, somatostatin and carbachol not only inhibit Ca^{2+} currents but also stimulate a 55 pS- K^+ channel in GH₃ cells²². On the basis of reconstitution

experiments with isolated membrane patches and purified or recombinant α subunits, the involvement of G_{i3} in K^+ -channel regulation has been proposed. Somatostatin also stimulates a 120 pS-, Ca^{2+} -activated K^+ channel³³; this effect is sensitive to pertussis toxin and may require a dephosphorylation step. Another example is found in atrial myocytes in which acetylcholine stimulates an inward rectifying K^+ channel; in cell-free systems, the effect of acetylcholine is mimicked by activated α subunits of all known G_i -like proteins³⁴. Our approach could help identify the G_i -type G proteins involved in hormonal regulation of different types of K^+ channels in pituitary cells or atrial myocytes. The method can be used principally to determine the function of sequenced Ca^{2+} -channel subunits³⁵, numerous G protein-coupled receptors^{36,37}, or subunits of receptor with intrinsic ion channel activity³⁸. Taken further, the method is suitable for determining the function of any protein, provided it can be measured in a single cell. □

Received 17 June; accepted 8 August 1991.

1. Birnbaumer, L. A. *Rev. Pharmac. Toxicol.* **30**, 675-705 (1990).
2. Simon, M. I., Strathmann, M. P. & Gautam, N. *Science* **252**, 802-808 (1991).
3. Graziano, M. P., Freissmuth, M. & Gilman, A. G. *J. Biol. Chem.* **264**, 409-418 (1989).
4. Mattera, et al. *J. Biol. Chem.* **264**, 465-471 (1989).
5. Asano, T., Semba, R., Kamiya, N., Ogasawara, N. & Kato, K. *J. Neurochem.* **50**, 1164-1169 (1988).
6. Hescheler, J., Rosenthal, W., Trautwein, W. & Schultz, G. *Nature* **325**, 445-447 (1987).
7. Schultz, G., Rosenthal, W., Trautwein, W. & Hescheler, J. A. *Rev. Physiol.* **52**, 275-292 (1990).
8. Offermanns, S. et al. *Molec. Endocr.* **5**, 995-1002 (1991).
9. Paulsen, E. J., Paulsen, R. H., Haugen, T. B., Gautvik, K. M. & Gørdeladze, J. O. *Molec. Cell. Endocr.* **76**, 45-53 (1991).
10. Mumby, S. M., Kahn, R. A., Manning, D. R. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **83**, 265-269 (1986).
11. Offermanns, S., Schultz, G. & Rosenthal, W. *J. Biol. Chem.* **266**, 3365-3368 (1991).
12. Rosenthal, W. et al. *EMBO J.* **7**, 1627-1633 (1988).
13. Cazenave, C. et al. *Nucleic Acids Res.* **17**, 4255-4272 (1989).
14. Zamecnik, P. C. & Stephenson, M. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 280-284 (1978).
15. Ewald, D. A., Sternweis, P. C. & Miller, R. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3633-3637 (1988).
16. Harris-Warrick, R. M. et al. *Neuron* **1**, 27-32 (1988).
17. Toselli, M., Lang, J., Costa, T. & Lux, H. D. *Pflügers Arch. ges. Physiol.* **415**, 255-261 (1989).
18. McFadzean, I., Mullaney, I., Brown, D. A. & Milligan, G. *Neuron* **3**, 177-182 (1989).
19. Tsukamoto, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2974-2978 (1991).
20. Bertrand, P., Sanford, J., Rudolph, U., Codina, J. & Birnbaumer, L. *J. Biol. Chem.* **265**, 18576-18580 (1990).
21. Strathmann, M., Wilke, T. M. & Simon, M. I. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6477-6481 (1990).
22. Birnbaumer, L., Abramowitz, J. & Brown, A. *Biochim. biophys. Acta* **1031**, 163-224 (1990).
23. Masters, S. E., Strout, R. N. & Bourne, H. R. *Protein Engng* **1**, 47-54 (1986).
24. Collu, R. et al. *Endocrinology* **122**, 1176-1178 (1988).
25. Silbert, S., Michel, T., Lee, R. & Neer, E. J. *J. Biol. Chem.* **265**, 3102-3105 (1990).
26. Izant, J. G. & Weintraub, H. *Cell* **36**, 1007-1015 (1984).

27. Weintraub, H., Izant, J. G. & Harland, R. M. *Trends Genet.* **1**, 22-25 (1985).
28. Green, P. J., Pines, O. & Inoué, M. A. *Rev. Biochem.* **55**, 569-597 (1986).
29. Shuttleworth, J. & Colman, A. *EMBO J.* **7**, 427-434 (1988).
30. Minshall, J. & Hunt, T. *Nucleic Acids Res.* **14**, 6433-6451 (1986).
31. Loke, S. L. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3474-3478 (1989).
32. Hélène, C. & Toulmé, J.-J. *Biochim. biophys. Acta* **1049**, 99-125 (1990).
33. White, R. E., Schonbrunn, A. & Armstrong, D. L. *Nature* **351**, 570-573 (1991).
34. Yatani, A. et al. *Nature* **336**, 680-682 (1988).
35. Mori, Y. et al. *Nature* **350**, 398-402 (1991).
36. Sunahara, R. K. et al. *Nature* **351**, 614-619 (1991).
37. Findlay, J. & Eliopoulos, E. *Trends Pharmac. Sci.* **11**, 492-499 (1990).
38. Changeux, J.-P. *Trends Pharmac. Sci.* **11**, 485-492 (1990).
39. Hsu, W. H. et al. *J. Biol. Chem.* **265**, 11220-11226 (1990).
40. Nomenclature Committee of the International Union of Biochemistry Eur. *J. Biochem.* **150**, 1-5 (1985).
41. Offermanns, S., Schultz, G. & Rosenthal, W. *Eur. J. Biochem.* **180**, 283-287 (1989).
42. Jackson, D. A., Yuan, J. & Cook, P. R. *J. Cell Sci.* **90**, 365-378 (1988).
43. Graessmann, M. & Graessmann, A. *Meth. Enzym.* **101**, 482-492 (1983).
44. Dreyer, C., Scholz, E. & Hausen, P. *Roux's Arch. dev. Biol.* **191**, 228-233 (1982).
45. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
46. Tanabe, T. et al. *Nature* **315**, 242-245 (1985).
47. Jones, D. T. & Reed, R. R. *J. Biol. Chem.* **262**, 14241-14249 (1987).

ACKNOWLEDGEMENTS. We thank C. Crane-Robinson (Portsmouth Polytechnic) for a critical discussion and revision of the manuscript; D. Wedlich (Institut für Molekularbiologie und Biochemie) for advice in immunochemical detection methods, and K. Spicher and K.-D. Hinsch (both of Institut für Pharmakologie) for antibody generation and purification. Our work was supported by grants from the Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie, and from the Trude-Goecke-Stiftung (B.W.).

LETTERS TO NATURE

Morphological differences between optical and infrared images of the spiral galaxy NGC309

David L. Block*† & Richard J. Wainscoat†

* Department of Computational and Applied Mathematics, Witwatersrand University, Private Bag 3, Wits 2050, South Africa
† Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

THE morphological classification of spiral galaxies into various types is suspected to be highly dependent on the wavelength of observation¹, because optical images emphasize young population I stars at the expense of other stellar types, as well as ionized gas and dust. Extension of the classification of galaxies out to wavelengths of a few micrometres has had to await the development of large-format near-infrared array cameras. We present here images at 2.1 μ m wavelength of NGC309, one of the largest 'grand design' (type ScI) spiral galaxies, obtained with the 256×256 array camera developed for the NICMOS (near-infrared camera

and multiobject spectrograph) instrument, designed for installation on the Hubble Space Telescope. Optically, NGC309 presents a classic multi-arm morphology, but at 2.1 μ m we see a two-arm spiral and the appearance of a prominent central bar; it resembles the SBa galaxy NGC1358. These studies underscore existing indications² that the disk structure of spiral galaxies may be unrelated to the Hubble type assigned from the transient population I morphology.

Pioneering work^{3,4} in the detection of smooth 'red' arms in the disk of the grand design spiral M51 (NGC5194) showed that the morphology of the evolved disk population need not mimic the Hubble classification assigned from the young population I tracers (OB stellar associations and ionized hydrogen regions). In 1976, Strom, Jensen and Strom⁵ commented: "Detailed study of the underlying density wave in late type spirals is considerably complicated by the dominance of the bright Population I tracers near the wave crests. This problem can be avoided, in part, by photographing these galaxies at near infrared wavelengths.... However, technical difficulties have thus far precluded such observations."

Previous investigations⁶⁻⁹ have been restricted to wavelengths shorter than 1 μ m. Here, we report galactic structure results at 2.1 μ m using a large format near-infrared array camera, for the grand-design Hubble type-c galaxy NGC309.

NGC309 is one of the most luminous spirals in the Revised

Shapley Ames Catalogue (RSA)¹⁰. At a redshift distance of 116 Mpc (assuming $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$), the absolute blue magnitude (corrected for foreground absorption and inclination) is -23.25 . The recession velocity has been confirmed using the 100-m radio dish at Effelsberg (W. Huchtmeier, personal communications). Sandage¹¹ describes NGC309 as being of the multiple-arm M101 type. The luminosity classification¹² is type I. There is an almost complete inner ring of young stars, so that the full classification given in the RSA is Sc(r)I.

Figure 1a shows the results of a photographically 'unsharp' masked and image-enhanced IIIa-J+GG 385 plate of NGC309, processed in the way described by Block¹³. The original plate was secured at prime focus of the 4-m Mayall telescope and the exposure time was 30 min. The photographic mask is needed to allow features of very low contrast, such as the bifurcation labelled at the end of one of the spiral arms, to be printed at the same time as the bright inner central regions (ring and bulge). Elmegreen and Elmegreen⁹ assign NGC309 to their arm classification 12 (extreme grand design). Messier 81, shown at the same absolute scale for comparison, is also assigned to an arm classification 12 and is of similar luminosity class (I-II).

Such photographic processing assists in appreciating the remarkable linear extent of this spiral ($\sim 100 \text{ kpc}$), highlighting

numerous segments or spurs¹⁴ that jut out from the spiral arms and may even extend from one spiral arm to the next. In the *Color Atlas of Galaxies*¹⁵ and the *Hubble Atlas*¹¹ the spiral arms appear narrow and filamentary, but the photographs in such atlases have not been scaled according to distance. When this is done¹⁶, the arms of NGC309, measured in kpc, are among the widest known. This can be seen in our Fig. 1a, and is predicted from the correlation between absolute magnitude and absolute arm width^{17,18}. So luminous is the galaxy that one supernova may be expected every two or three years^{19,20}. We now examine NGC309 at near-infrared wavelengths (of $2.1 \mu\text{m}$ and less) to explore morphological differences between Fig. 1a and the infrared images.

The University of Hawaii's 256×256 near-infrared camera has only recently been commissioned and made available for general use. The camera uses a Rockwell HgCdTe array, sensitive from 1 to $2.5 \mu\text{m}$. At the $f/10$ focus of the university's 2.2-m telescope, with 1:1 reimaging optics, the scale is 0.37 arcsec per pixel, giving a field of $95 \times 95 \text{ arcsec}$. The filter used was K' (central wavelength $2.1 \mu\text{m}$); it is similar to K, but shifted $0.1 \mu\text{m}$ towards the blue to reduce the thermal background. The filter will be discussed further elsewhere (R.J.W. and L. Cowie, manuscript in preparation).

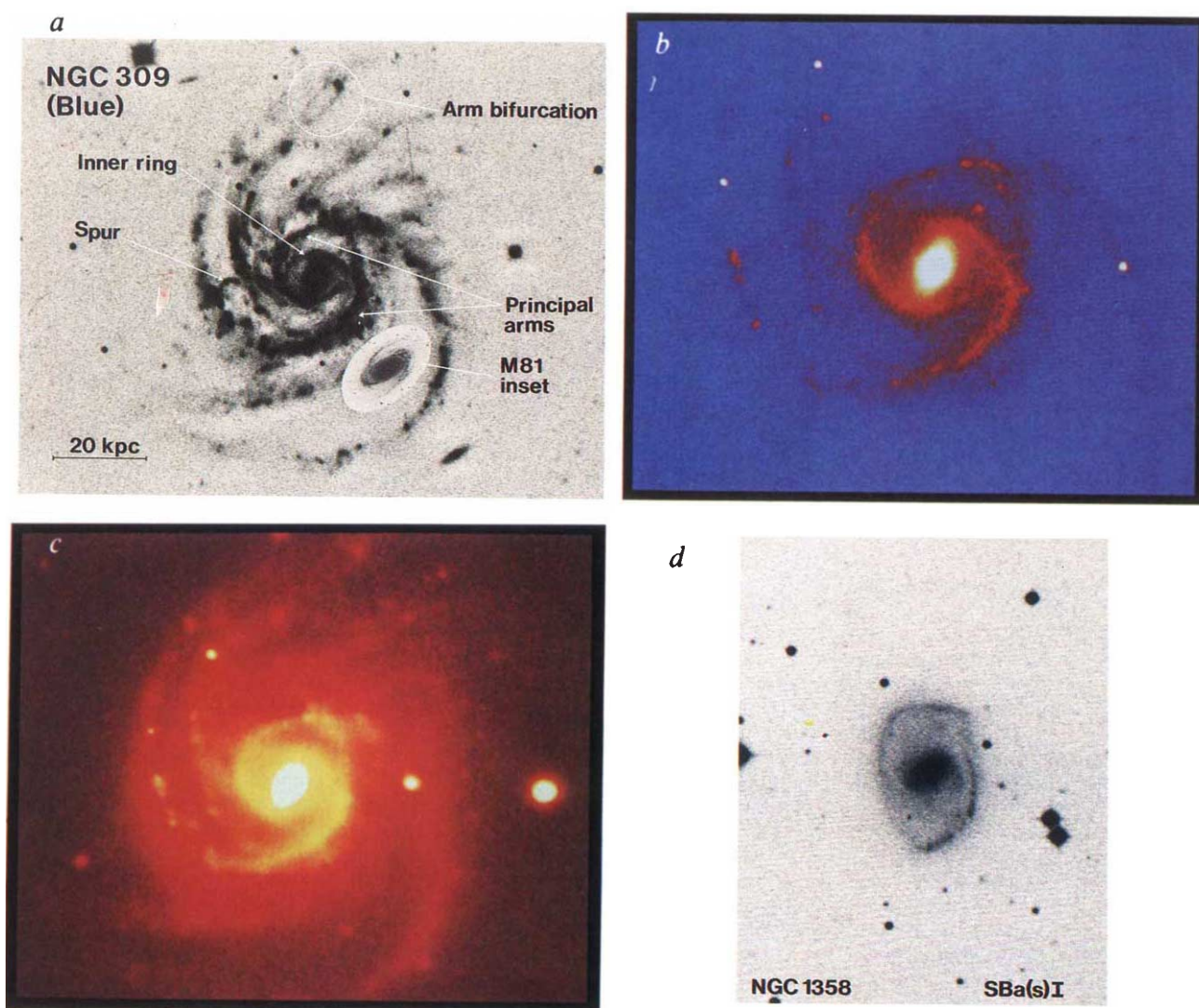


FIG. 1 a, A photographically unsharp masked and image-enhanced IIIa-J plate of NGC309; note inset of extreme grand design spiral Messier 81, for comparison. The 20-kpc scale bar applies to both the image of NGC309 and the inset. b, $2.1\text{-}\mu\text{m}$ image of NGC309 secured at Mauna Kea using the new 256×256 infrared array. c, Palomar 5-m image of NGC309 with Gunn

four-shooter (i filter, $0.83 \mu\text{m}$). Orientation for all the NGC309 images is north up, east to the left. d, Equatorial J-Survey photograph of the SBa spiral NGC1358, secured with the UK Schmidt Telescope, courtesy of the Royal Observatory (Edinburgh).

The near-infrared imaging of NGC309 was performed in four quadrants, the bulge being common to each of the two southern quadrants. Six frames, shifted slightly relative to each other, with exposure times of 90 s, were added for each quadrant, and the four quadrants added to form the final image shown in Fig. 1b, which reveals the principal material arms. Separate sky observations were made for sky subtraction, and dome flats used for flat-fielding. Surface photometry of NGC309 at 2.1 μm will be presented elsewhere, but we do wish to mention that at K' the luminosity profile of the disk is exponential, with a disk scale length of 18.34 arcsec or 10.3 kpc (at our assumed distance of 116 Mpc).

While we conducted our 2.1- μm observations at Mauna Kea, the galaxy was also observed for us at shorter wavelengths using the g and i filters on the Gunn four-shooter attached to the 5-m Hale reflector. Each of the four CCD chips covers 800×812 pixels, with a scale of 0.33 arcsec per pixel. Images obtained through the Gunn g and i filters roughly correspond to the V (0.54 μm) and I (0.83 μm) passbands respectively. Exposure times were 60 s in both g and i; longer exposures (300 s each) were also taken. Figure 1c shows one of the Palomar 60-s images through the Gunn i filter. The underlying disk is colour-coded red, the highest surface brightness features (density wave crests) are coded yellow. Unlike flocculent spirals, which show very little structure in their old stellar population disks, extreme grand design spirals (such as NGC309) have rich disk structure.

It is instructive to note how the overall morphology changes as one probes the disk features, rather than the population I features. The inner (r) ring of young stars is no longer dominant; in fact the morphology changes from ring-shaped (r) to s-shaped (s). Furthermore, instead of the three-arm appearance seen in B (IIIa-J), only two short principal spiral arms are seen at K' (Fig. 1b). The spiral becomes a two-arm grand design, with a prominent bar, and now bears a striking resemblance to the SBa galaxy NGC1358 (Fig. 1d). Comparing Fig. 1b with d, it is at first difficult to believe that one is not looking at the same galaxy.

The fractions of the energy radiated by a synthetic spiral galaxy in the different filter bands U, B, V, I, K and L are presented in Fig. 1 of Johnson²¹; at $\sim 2 \mu\text{m}$, one is essentially probing the evolved K and M giants. The great advantage is that one is able to trace the distribution of mass in the disk. The mass distribution for NGC309 is radically different from its Hubble type, confirming the findings of Burstein and Rubin² that disk dynamics and Hubble type are uncorrelated.

The two material arms, which show up easily and unambiguously at $\sim 2 \mu\text{m}$, are characterized by a $65 \pm 10\%$ surface brightness enhancement compared with the underlying disk. The quoted uncertainty comes principally from the photon noise (which makes the underlying light level and the peak level uncertain) and also includes uncertainties in sky subtraction down to a surface brightness of $K' = 22 \text{ mag arcsec}^{-2}$. The arms lie within a radius of 17 kpc from the galaxy centre. The corresponding position of the arms in the blue image are indicated in Fig. 1a. These two smooth arms bear a striking resemblance to the 'red' arms discovered by Zwicky³ on composite photographs of M51. Extreme grand design spirals (such as M51) have strong near-infrared arms, with amplitudes $\sim 60\%$ above the background disk light^{7,9}.

The degree of development of spiral arms (van den Bergh luminosity class¹²) clearly depends very much on bandpass, especially as far as winding angle is concerned. Whereas the winding angles of the population I arms in Fig. 1a are large, the surface brightness of the two principal arms in our 2.1- μm image decreases abruptly after a winding angle of only $\sim 180^\circ$. A very similar effect is seen for M51. On blue emulsions, the winding angle for the population I tracers in the longest arm²² is 490° , whereas the smooth red arms curl around just one-half revolution before their surface brightness drops abruptly.

Disk morphology is not transient: the density wave patterns may persist for the entire Hubble time, unlike the relatively

short-lived population I tracers. According to the density wave theory for NGC309, it would be only two arms in the disk, coupled with a prominent bar, that are responsible for 'driving the wave', producing the necessary shock fronts and compressions to give rise to the multiple arms seen at shorter wavelengths.

The point we make in extending the classification procedure out to a few micrometres is that if one were presented with the K' image of NGC309 and asked to classify the galaxy from its mass distribution, the winding and pitch angles of the two short arms, the overall 'smoothness' of those arms at 2.1 μm , the presence of the prominent bar and the disappearance of the inner ring would suggest a classification very different from that assigned at shorter wavelengths. Large-format infrared arrays, operative well beyond I, now open a new era in the classification of disks, probing material arms responsible for the transient population I morphology, on which Hubble's original scheme is largely based. $\square$

Received 16 May; accepted 16 July 1991.

- Bothun, G. D. Conf. Ser. astr. Soc. Paci. **10** (ed. Kron, R. G.) 54-66 (1990).
- Burstein, D. & Rubin, V. C. *Astrophys. J.* **297**, 423-435 (1985).
- Zwicky, F. *Publ. astr. Soc. Pacif.* **67**, 232-236 (1955).
- Zwicky, F. *Morphological Astronomy* 194-204 (Springer, Berlin, 1957).
- Strom, S. E., Jensen, E. B. & Strom, K. M. *Astrophys. J.* **206**, L11-L14 (1976).
- Sharpless, S. & Franz, O. G. *Publ. astr. Soc. Pacif.* **75**, 219-225 (1963).
- Schweizer, F. *Astrophys. J. Suppl.* **31**, 313-332 (1976).
- Elmegreen, D. M. *Astrophys. J. Suppl.* **47**, 229-233 (1981).
- Elmegreen, D. M. & Elmegreen, B. G. *Astrophys. J. Suppl.* **54**, 127-149 (1984).
- Sandage, A. & Tammann, G. A. *A Revised Shapley-Ames Catalogue of Bright Galaxies* (Carnegie Institution, Washington DC, 1981).
- Sandage, A. *Hubble Atlas of Galaxies* (Carnegie Institution, Washington DC, 1961).
- van den Bergh, S. *Astrophys. J.* **131**, 215-223 (1960).
- Block, D. L. *Nature* **347**, 452-455 (1990).
- Weaver, H. F. *Interstellar Gas Dynamics* (ed. Habing, H. J.) IAU Symp. No. 39, 22-50 (Reidel, Dordrecht, 1970).
- Wray, J. D. *Color Atlas of Galaxies* (Cambridge University Press, 1988).
- Block, D. L. *A Photographic Atlas of Primarily Late Type Spirals Printed as if Each Galaxy were at the Same Distance* (University of Fort Hare Press, Ciskei, 1984).
- Block, D. L. *Astr. Astrophys.* **109**, 336-339 (1982).
- Kennicutt, R. C. *Astr. J.* **87**, 255-263 (1982).
- Zwicky, F. in *Stars and Stellar Systems* **8**, ch. 7 (ed. Aller, L. H. and McLaughlin, D. B.) (University of Chicago Press, 1965).
- Tammann, G. A. *Astr. Astrophys.* **8**, 458-475 (1970).
- Johnson, H. L. *Astrophys. J.* **143**, 187-191 (1966).
- Danver, C. G. *Lund Obs. Ann.* **10** (Lund Observatory, 1942).

ACKNOWLEDGEMENTS. D.L.B. thanks D. Hall and A. Stockton for hospitality and use of facilities at the Institute for Astronomy, Hawaii. We thank S. Wolff for granting 4-m telescope time at NOAO, W. Freedman for Gunn 4-shooter data, T. Kinman and S. Tritton for optical data, L. Cowie, J. Kormendy, J. Cromer, W. Huchtmeier, W. Keel, J. Wray, R. West, T. Farnham, J. Elias, G. Jacoby, C. Madsen and B. Dumoulin for their input and H. van der Laan, F. Sellschop, R. Arndt and P. Tyson for support. The 2.1- μm observations were obtained at the University of Hawaii's 2.2-m telescope with the University of Hawaii Infrared Camera, which uses a $256 \times 256 \text{ HgCdTe}$ array developed by Rockwell Science Center.

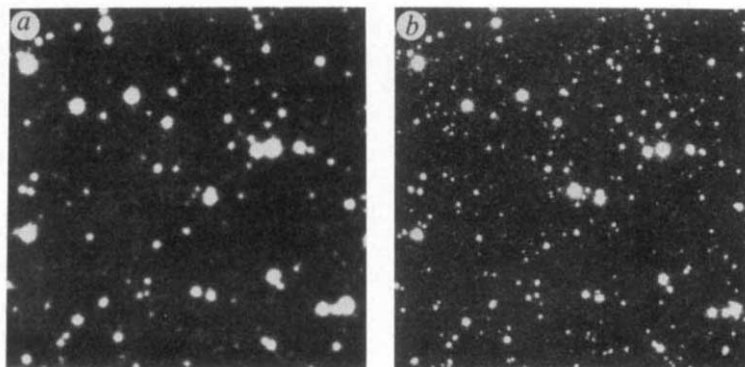
Evidence for a black hole in the X-ray nova Muscae 1991

M. Della Valle, B. J. Jarvis & R. M. West

European Southern Observatory, Casilla 19001, Santiago 19, Chile

X-RAY novae form a subclass of low-mass X-ray binaries, which typically consist of a neutron star accreting material from a low-mass, late-type companion. In a few cases, however—notably A0620-00 (ref. 1)—, the accreting object may be a black hole. The detection and study of the optical counterpart is then of great importance in understanding the origin of the X-ray emission and the nature of the companion. We report here the discovery of a new X-ray nova, Nova Muscae 1991, which we found in a search for the optical counterpart of the new transient X-ray source GRS1121-68. The optical position, light-curve and early spectral development of the outburst support the identification of Nova Muscae 1991 as a binary, and similarities with A0620-00 (known optically as V616 Mon) in its early evolutionary stages led us to classify Nova Muscae 1991 as a candidate black-hole binary.

FIG. 1 *a*, Reproduction of an earlier red-sensitive ESO Schmidt plate (120-min exposure on IIIa-F+RG630; 29 January 1976). A total of twelve astrometric standards on twelve photographic plates covering the area around Nova Muscae 1991 were used to determine the position of the possible pre-nova seen as a faint, star-like object on the deepest pre-outburst plates. The resultant mean difference in position between the pre-nova star and the nova was $0.32''$, which is within the astrometric uncertainties and which confirms the identification of the pre-nova of Nova Muscae 1991. Pre-nova magnitudes measured from these plates yield $B=20.9\pm0.2$ and $R=19.4\pm0.1$. *b*, Nova Muscae 1991 during its flare up in early January 1991 in the southern constellation of Musca. The CCD frame was secured on 15.3 January 1991 (seeing 0.9 arcsec) with the ESO New Technology Telescope (5-s exposure in R). The nova is the bright object at the centre.



The X-ray transient source GRS1121-68 was almost simultaneously detected by the All Sky X-ray Monitor aboard Ginga on 8 January 1991 and by the *Watch* all-sky X-ray camera on the Soviet GRANAT satellite on 9 January². The object was identified³ as a star of $V\approx13.5$ on a IIIa-J plate taken on 13 January with the 1-m Schmidt telescope, near the edge of the error box reported by Makino and the Ginga team⁴ (Fig. 1*a*) and later confirmed by X-ray and γ -ray observations⁵. Photometric and spectroscopic observations were immediately carried out with the New Technology Telescope (NTT) and the 2.2-m and 1.5-m telescopes at La Silla Observatory^{6,7}. At the position of the nova, RA = 11 h 24 min 18.49 s, Dec. = $-68^\circ 24' 01.7''$ (1950.0), the progenitor is barely visible on the best ESO (R) Schmidt plates (1984) with a magnitude close to $R\approx20$ (Fig. 1*b*).

The striking feature of the photometric evolution of Nova Muscae 1991 was the large increase of brightness (about 8 magnitudes) which led to a classification of the object, in the first instance, as a classical nova. The epoch of maximum light (see Fig. 2) probably took place between 11 January (Julian day JD = 2,448,268) and 12 January (JD = 2,448,269) at a magnitude $V\approx13.3$ and $B\approx13.5$. We caution, however, that this estimate may be somewhat inaccurate because it is based on photographic magnitudes only and may also suffer from fluctuations as seen at later stages (see also ref. 8).

The optical light curve of Nova Muscae 1991 closely resembles the optical evolution of V616 Mon (ref. 9) in the amplitude of the outburst, the decay rate (~0.024 mag day⁻¹) calculated over the first 45 days past maximum and the variations in brightness (≈0.2 mag) on a timescale of ~1 day, observed in the three photometric bands. This decline seems slower than that shown by Cen X-4, which faded by two magnitudes in the first 30 days past maximum.

The spectrum (Fig. 3) obtained with the NTT telescope on 15 January (resolution ~10 Å) does not resemble that of a classical nova at an early stage. We note the complete absence of Fe II, Na I, O I and Balmer emission lines coupled with *P* Cyg profiles which normally predominate in the pre-maximum and principal spectra of novae at early stages¹⁰. The spectrum of Nova Muscae 1991 is dominated by Balmer, He and N

emission lines superimposed on a blue continuum. Comparison with other well studied X-ray novae shows that this spectrum closely resembles that of the optical counterpart of Cen X-4 (ref. 11) and V616 Mon (refs 12-14).

Table 1 summarizes the main spectroscopic features visible in the spectrum of Fig. 2. First, there is the N III blend at 4,640 Å equivalent width (EW) = 3.5 Å. The N III emission is normally attributed to the X-ray heating driving the Bowen fluorescence process¹⁵, but in the present case it appears broadened by the O II and C III ions. On the red wing we see the He II line at 4,686 Å (EW = 2.4 Å). Second, the Balmer lines are seen in emission filling in shallow absorptions. Values for the Balmer decrement as deduced from the table are $H\alpha/H\beta\approx4$ and $H\gamma/H\beta\approx0.5$. Third, the main interstellar features are the blended (EW = 0.9 Å) interstellar bands at 5,780, 5,778 and 5,797 Å, the unresolved Na D lines (EW = 1.4 Å) and the blend due to the absorption lines at 6,269 and 6,282 Å (EW = 0.5 Å).

We can estimate the colour excess several ways. A preliminary indication comes from the strength of the interstellar Na D absorption lines. Using the empirical relation recently revised by Barbon *et al.*¹⁶ between the colour excess and the width of the Na D lines, and taking EW = 1.4 Å, we estimate a colour excess of $E(B-V)=0.3$. This is consistent both with the IUE measurement¹⁷ of $E(B-V)=0.20\pm0.05$, arising from the dust absorption at 2,200 Å, and with the value of $E(B-V)=0.35$ which we have derived from the strength of the absorption bands¹⁸ at 5,778, 5,780 and 5,797 Å. In the following we shall adopt $\langle E(B-V) \rangle = 0.30\pm0.10$.

The estimate of the distance of Nova Muscae is rather uncertain. As Nova Muscae 1991 is not a classical nova, the relation between the absolute magnitude at maximum and the rate of decline, which is well established for galactic novae, cannot be successfully applied. The relation derived by Capaccioli *et al.*¹⁹

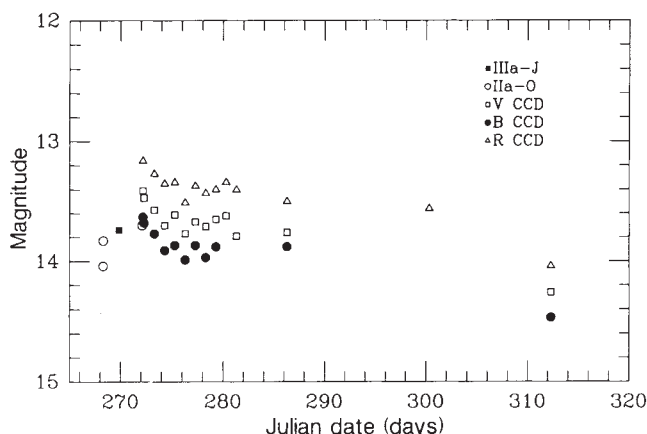


FIG. 2 Light curve of Nova Muscae 1991 during the early stages after outburst in the B, V and R colours.

TABLE 1 Observed emission lines

Line	$\lambda(\text{\AA})$	Flux (10^{-16} erg s ⁻¹ cm ⁻²)
H γ	4,340	150
N III	4,640	800
He II	4,686	550
H β	4,861	300
He I	5,876	100
H α	6,563	1,200
He I + He II	6,678-6,683	81
He I	7,068	85
N II	7,220	70

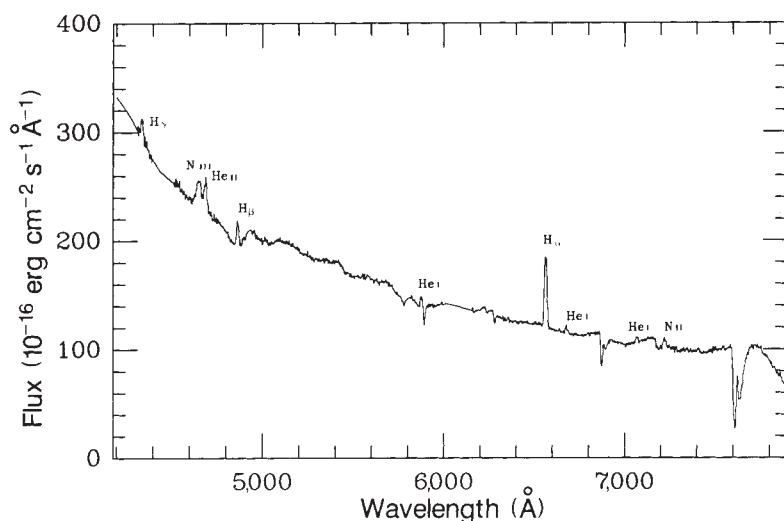


FIG. 3 Spectrum of Nova Muscae 1991 in the range 4,000–8,000 Å (resolution ~ 10 Å), obtained on 15 January 15.3 ut at NTT.

for galactic novae, combined with the values derived above ($V_{\max} = 13.3$, $E(B - V) = 0.30$, rate of decline $V_d = 0.024 \text{ mag d}^{-1}$), yields a distance of $d \approx 60 \text{ kpc}$, which is very unlikely. Alternatively, we can estimate the distance from the linear relation between the equivalent width of the Na D line and distance²⁰. In this case we get the more reasonable value of $d \approx 1.4 \text{ kpc}$. Combining this distance with measurements of the integrated flux between 1,200 and 3,200 Å ($1.1 \times 10^{-10} \text{ erg cm}^{-2} \text{ s}^{-1}$) (ref. 14), between 4,000 and 8,000 Å ($\sim 0.7 \times 10^{-10} \text{ erg cm}^{-2} \text{ s}^{-1}$), and between 35–150 keV ($\sim 2 \times 10^{-8} \text{ erg cm}^{-2} \text{ s}^{-1}$) (ref. 3), we get an ultraviolet plus optical luminosity at maximum of $L_{\text{opt}}^{\max} \approx 4 \times 10^{34} \text{ erg s}^{-1}$, an X-ray luminosity of $L_X^{\max} \approx 10^{37} \text{ erg s}^{-1}$, and a ratio $L_X/L_{\text{opt}} \geq 100$. These values are typical for low-mass X-ray binary stars in outburst²¹.

Although the photometric evolution of Nova Muscae 1991 resembles that of a classical nova in some respects, both the spectral features and the amount of the energy released during the outburst fit the characteristics of the optical counterparts of X-ray sources such as Sco X-1, Cen X-4 and V616 Mon. With $d = 1.4 \text{ kpc}$ and $B = 19.8$ (magnitude at minimum after reddening correction), we derive a magnitude $M_B \approx +9$ for the progenitor. This is consistent with the absolute magnitude of a low-mass main-sequence star of spectral type $\sim K5\text{--}M0$ (ref. 22). According to current understanding, these late-type stars are the most probable companions of the degenerate stars in low-mass X-ray binary systems²¹.

These observational data support the idea that the energy source during the outburst was the gravitational potential energy released during the transfer of matter from the accretion disk onto the compact companion. This energetic production, according to current models²¹, can be triggered by a mass-transfer rate from the accretion disk onto the compact object smaller than that from the secondary onto the accretion disk (disk-instability model), and/or through a sudden increase in the mass-transfer rate from the secondary onto the accreting object (mass-transfer instability model). Finally, the large amount of energy released at X-ray wavelengths, at least $10^3\text{--}10^4$ times as large as that emitted by a normal cataclysmic variable, indicates that the compact object that has accreted material from the K-M dwarf is probably a neutron star, but does not exclude a black hole^{1,23}.

The binary nature of Nova Muscae should be observable during the next few months, when it is expected that the star will revert to minimum light. □

4. Makino, F. *et al.* *IAU Circ.* No. 5161 (1991).
5. Sunyaev, R., Jourdain, E. & Laurent, P. *IAU Circ.* No. 5176 (1991).
6. Della Valle, M. & Pakull, M. *IAU Circ.* No. 5167 (1991).
7. Della Valle, M., Jarvis, B. J. & West, R. *ESO Messenger* **63**, 3 (1991).
8. Bailyn, C. *IAU Circ.* No. 5259 (1991).
9. Lloyd, C., Nögle, R. & Penston, M. V. *Mon. Not. R. astr. Soc.* **179**, 675 (1977).
10. Payne-Gaposchkin, C. in *The Galactic Novae*, ch. 10 (1957).
11. Canizares, C. R., McClintock, J. E. & Grindlay, J. *Astrophys. J.* **236**, L55 (1980).
12. Whelan, J. A. *et al.* *Mon. Not. R. astr. Soc.* **180**, 657 (1977).
13. Boley, F. *et al.* *Astrophys. J.* **203**, L13 (1976).
14. Ciatti, F., Mammano, A. & Vittone, A. *Astr. Astrophys.* **56**, 311 (1977).
15. McClintock, J. E., Canizares, C. R. & Tarter, C. B. *Astrophys. J.* **198**, 641 (1975).
16. Barbon, R., Benetti, S., Cappellaro, E., Rosino, L. & Turatto, M. *Astr. Astrophys.* **237**, 79 (1990).
17. Gonzales-Riestra, R. *IAU Circ.* No. 5174 (1991).
18. Herbig, G. H. *Astrophys. J.* **196**, 129 (1975).
19. Capaccioli, M., Della Valle, M., D'Onofrio, M. & Rosino, L. *Astrophys. J.* **97**, 1622 (1989).
20. Charles, P. A. *Proc. 23rd ESLAB Symp.* **103** (European Space Agency, Paris, 1989).
21. Van Paradijs, J. & Verbunt, F. in *High Energy Transient in Astrophysics* (ed. Woosley, S. E.) **49** (1984).
22. Allen, C. W. in *Astrophysical Quantities*, **206** (1976).
23. White, N. E. *Astr. Astrophys. Rev.* **1**, 85 (1989).

Extension of Euler's theorem to symmetry properties of polyhedra

A. Ceulemans* & P. W. Fowler†

* Department of Chemistry, University of Leuven, Celestijnenlaan 200F, B-3001 Leuven, Belgium

† Department of Chemistry, University of Exeter, Stocker Road, Exeter EX4 4QD, UK

POLYHEDRAL cages and clusters are widespread in chemistry. Examples of fully triangulated polyhedra (deltahedra) are the skeletons of *clos*-boranes $B_n H_n^{2-}$, many heteroboranes and transition-metal carbonyls¹. Three-connected cages occur for carbon^{2,3} and in zeolites¹. The numbers v , f and e of vertices, faces and edges of a convex polyhedron are related by Euler's theorem^{4,5} $v + f = e + 2$. Here we show that within the point group of the polyhedron the symmetries spanned by the sets of vertices, faces and edges are also related. We prove a general theorem relating these symmetries for convex polyhedra, and give further relations specific to deltahedra and 3-connected polyhedra. The latter extensions of Euler's theorem to point-group characters allow us to generate complete sets of internal vibrational coordinates from bond stretches for deltahedra, and to classify, from symmetry properties alone, the bonding or antibonding nature of molecular orbitals of 3-connected cages.

Our first relationship is an extension of the general Euler's theorem to group-theoretic representations Γ . Recall that for a pseudospherical cage it is useful to define σ , π , δ ... representations for sets of equivalent positions (or 'orbits') by considering their relation to a central point unshifted by all operations of

Received 9 May; accepted 16 July 1991.

1. McClintock, J. E. & Remillard, R. A. *Astrophys. J.* **308**, 110 (1986).

2. Lundt, N. & Brandt, S. *IAU Circ.* No. 5161 (1991).

3. Della Valle, M., Jarvis, B. J. & West, R. *IAU Circ.* No. 5165 (1991).

TABLE 1 Symmetry properties of the deltahedral skeletons of *closa*-boranes

Atoms	Polyhedron	Point group	Vertices*	Edges	$\Gamma_\sigma(e)$	$N(k)^*$
4	Tetrahedron	T_d	O_4	O_6	$A_1 + E + T_2$	3
5	Trigonal bipyramid	D_{3h}	$O_2 + O_3$	$O_3 + O_{6v}$	$2A'_1 + 2E' + A'_2 + E''$	8
6	Octahedron	O_h	O_6	O_{12}	$A_{1g} + E_g + T_{2g} + T_{1u} + T_{2u}$	5
7	Pentagonal bipyramid	D_{5h}	$O_2 + O_5$	$O_5 + O_{10v}$	$2A'_1 + 2E'_1 + 2E'_2 + A'_2 + E''_1 + E''_2$	12
8	Triangular dodecahedron	D_{2d}	$2O_{4d}$	$O_2 + O_4 + O_{4d} + O_8$	$4A_1 + A_2 + 2B_1 + 3B_2 + 4E$	30
9	Tricapped trigonal prism	D_{3h}	$O_3 + O_{6v}$	$O_3 + O_{6v} + O_{12}$	$3A'_1 + A'_2 + 4E' + A''_1 + 2A''_2 + 3E''$	27
10	Bicapped square antiprism	D_{4d}	$O_2 + O_{8d}$	$O_8 + 2O_{8d}$	$3A_1 + B_1 + 2B_2 + 3E_1 + 3E_2 + 3E_3$	28
11	Octadecahedron	C_{2v}	$O_1 + 2O_{2xz} + O_{2yz} + O_4$	$O_1 + 2O_{2xz} + O_{2yz} + 5O_4$	$9A_1 + 5A_2 + 7B_1 + 6B_2$	109
12	Icosahedron	I_h	O_{12}	O_{30}	$A_g + G_g + 2H_g + T_{1u} + T_{2u} + G_u + H_u$	9

* O_n is an orbit⁶ (a set of equivalent points) in the symmetry group and $N(k)$ is the number of independent harmonic force constants for the cage.

the point group⁶. Structureless points, radial vectors and s orbitals span Γ_σ ; pairs of tangential vectors or p orbitals span Γ_π . We can define σ representations for face centres, edge centres and vertices of a polyhedron. The new symmetry theorem is then

$$\Gamma_\sigma(v) \times \Gamma_\sigma(f) + \Gamma_\sigma(f) = \Gamma_\perp(e) + \Gamma_0 + \Gamma_e \quad (1)$$

where Γ is a representation (reducible or irreducible), f , v and e stand for faces, vertices and edges and the subscript $\perp$ denotes tangential vectors, one perpendicular to each edge. Γ_0 is the totally symmetric representation and Γ_e is the pseudoscalar representation, which has character +1 for proper and -1 for improper (parity reversing) operations. For the identity operation nothing is shifted and (1) reduces to the familiar $v + f = e + 2$; more generally it implies that the symmetry of any one set of structural components is related to the other two.

An elementary proof of (1) follows from consideration of the characters $\chi_\sigma(v)$, $\chi_\sigma(f)$, $\chi_\perp(e)$ under point-group operations. Symmetry elements of a convex polyhedron can intersect its surface in a limited number of ways. In each hemisphere a rotational axis C_n must pass through a vertex, a face centre or (for C_2 only) an edge centre. Each vertex on the axis adds 1 to $\chi_\sigma(v)$, each face adds 1 to $\chi_\sigma(f)$ and each edge on a C_2 axis -1 to $\chi_\perp(e)$; the sum $\chi_0 + \chi_e$ is +2. For all six combinations ($v + v$, $v + e$, ..., $f + f$), therefore, (1) is satisfied. For an improper operation $\chi_0 + \chi_e$ vanishes, and as all structural components are shifted by inversion and rotation-reflections, $\chi_\sigma(v) = \chi_\sigma(f) = \chi_\perp(e) = 0$ so that (1) is trivially satisfied for these operations. Pure reflections remain to be considered. Edges intersecting a

mirror plane lie either wholly within the plane or at right angles to it, and the mirror must therefore cut the polyhedron in a closed chain made up of links of at most four types. Figure 1 shows that each link contributes equally to both sides of (1), and hence the equation holds for all symmetry elements.

An alternative form of (1) follows by considering duals. Every polyhedron P has a face-dual D defined by the fact that the vertices of D lie at the face centres of P and face centres of D at the vertices of P ; edges of D cross those of P at right angles. Noting that (1) applies to D , that $\Gamma_\sigma(v)$ for D is $\Gamma_\sigma(f)$ for P and that $\Gamma_\perp(e)$ for D is $\Gamma_\parallel(e)$ for P , it must be true for all convex polyhedra that

$$\Gamma_\sigma(f) \times \Gamma_\sigma(v) + \Gamma_\sigma(v) = \Gamma_\parallel(e) + \Gamma_0 + \Gamma_e \quad (2)$$

The same result could be obtained by multiplying (1) with Γ_e . Equation (2) is a second statement of the generalized Euler theorem, referring now to vectors along the edges. This rotation of edge vectors through 90° precisely corresponds to the parity transformation of p orbitals in Stone's tensor surface harmonic (TSH) theory of the electronic structure of cluster orbitals^{7,8}. As $\Gamma_\perp(e)$ and $\Gamma_\parallel(e)$ form disjoint halves of the π representation of the edges, (1) and (2) add to give

$$\begin{aligned} [\Gamma_\sigma(f) + \Gamma_\sigma(v)] \times (\Gamma_0 + \Gamma_e) &= 2(\Gamma_0 + \Gamma_e) + \Gamma_\pi(e) \\ &= 2(\Gamma_0 + \Gamma_e) + \Gamma_\sigma(e) \times \Gamma_T - \Gamma_\sigma(e) \end{aligned} \quad (3)$$

where Γ_T is the translational representation and the second equality follows from the general relation⁹: $\Gamma_\pi = \Gamma_\sigma \times \Gamma_T - \Gamma_\sigma$. Equation (3) shows that permutational representations of all three types of structural component are interdependent, as are their numbers in the original Euler theorem.

As a specific example of equations (1)–(3), consider the cube, in O_h symmetry $\Gamma_0 = A_{1g}$ and $\Gamma_e = A_{1u}$. The vertices span $A_{1g} + T_{2g} + A_{2u} + T_{1u}$ and the faces $A_{1g} + E_g + T_{1u}$. Therefore (1) predicts that 12 vectors or p orbitals across the edges span $A_{2g} + E_g + T_{1g} + T_{1u} + T_{2u}$, and equation (2) predicts that the 12 vectors along the edges span $A_{2u} + E_u + T_{1u} + T_{1g} + T_{2g}$. Equation (3) implies that the sum of these two reducible representations should be Γ_π for the 12-orbit of O_h . All three predictions are easily verified by direct inspection.

A physical interpretation can be given in terms of fluid flow on the surface of a polyhedron. Each theorem expresses the symmetry of a continuity equation. On the left-hand side of (1), $\Gamma_\sigma(f)$ represents the density that builds up at face centres, whilst $\Gamma_e \times \Gamma_\sigma(v)$ corresponds to the symmetry of rotations around axes through the vertices. The $\Gamma_\perp(e)$ term represents flow across the edges. Equation (1) then describes a conservation property: changes of density at faces and vertices will match the flow across the edges except for the scalar term, which corresponds to a uniform density change at all face centres, and the pseudoscalar term which represents a simultaneous rotation around all vertices; neither causes any flow across the edges. Equation (2) has a similar interpretation in terms of rotation about faces, density changes at vertices and flow along edges.

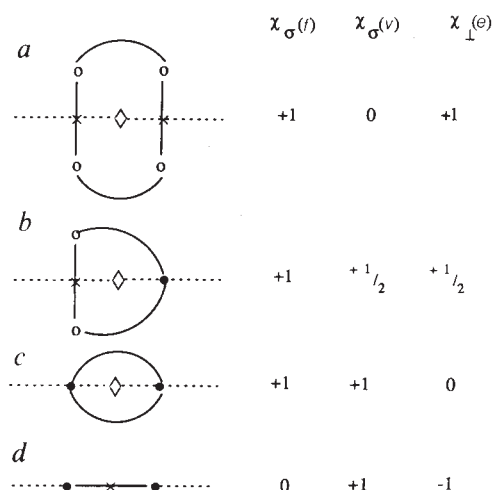


FIG. 1 Intersection of a reflection plane with the surface of a polyhedron. Linked units in the chain may be a face with (a) two, (b) one or (c) no perpendicular edges, or (d) an in-plane edge. To avoid double-counting the contribution to the character must be halved for shared elements (vertices $\bullet$, edge centres $\times$) but counted fully for the unshared elements (face centres $\diamond$) of a unit.

From general convex polyhedra we turn now to deltahedra. The starting point for our derivation of another analogue of Euler's theorem was a remark by Boyle and Parker¹⁰ to the effect that the bond stretches (that is, the edges) span the same representation as the internal vibrations for an icosahedral cage. It turns out that edge and vibrational representations are equivalent for all convex deltahedra; that is,

$$(\text{deltahedron}) \quad \Gamma_{\sigma}(v) \times \Gamma_T - \Gamma_T - \Gamma_R = \Gamma_{\sigma}(e) \quad (4)$$

where Γ_R is the rotational representation, $\Gamma_T \times \Gamma_e$.

Proof of (4) by character arithmetic follows the same lines as for (1). Under the identity (4) follows from dimensional considerations. For a convex solid with solely triangular faces, C_n axes may intersect edges and vertices ($n=2$), face centres and vertices ($n=3$), or vertices ($n \geq 4$). χ_T and χ_R both have the value $1 + 2 \cos(2\pi/n)$ and consideration of cases proves (4) for all C_n . For the reflection plane $\chi_T + \chi_R = 0$ and the only allowed links in the chain are now two-triangle kites (Fig. 2), each contributing +1 to $\chi_{\sigma}(e)$ and $\chi_{\sigma}(v)$, so that again (4) holds.

Table 1 gives a symmetry analysis for deltahedra with 4 to 20 faces and lists the sets of equivalent vertices and edges as 'orbits'. An immediate consequence of (4) is that a general vibrational force field for a deltahedral skeleton can be couched entirely in terms of stretching coordinates. If the harmonic potential energy is

$$V = \frac{1}{2} \sum k_{ij} r_i r_j$$

there are $e(e+1)/2$ permutationally distinct terms but only $N(k)$ independent force constants. $N(k)$ is equal to the number of copies of Γ_0 in the symmetric square of the vibrational representation¹¹ or, by (4), of the edge representation (Table 1). Physically, (4) expresses the mechanical stability of triangulated structures¹². If a deltahedron had completely rigid bonds it could not deform; the structure has no freedom to bend or twist except through changes in bond length.

Face and edge representations for convex deltahedra are also related. One equation easily proved with characters is

$$(\text{deltahedron}) \quad \Gamma_{\sigma}(f) \times \Gamma_T = \Gamma_{\perp}(e) + \Gamma_{\sigma}(e) \quad (5)$$

For a deltahedron, application of Euler's theorem leads to $3v - 6 = e$ and $3f = 2e$. Theorem (4) is a character extension of the first of these dimensionality relations, theorem (5) of the second. We note that any convex polyhedron is converted to a deltahedron by capping all faces. Capping a deltahedron itself produces matching changes to $\Gamma_{\sigma}(v) \times \Gamma_T$ and $\Gamma_{\sigma}(e)$ because each cap adds one vertex but three edges that span the local cartesian representation. Multiplying (5) by Γ_e gives

$$(\text{deltahedron}) \quad \Gamma_{\sigma}(f) \times \Gamma_R = \Gamma_{\parallel}(e) + \Gamma_{\sigma}(e) \times \Gamma_e \quad (6)$$

Addition of (5) and (6) and use of the definition of $\Gamma_{\pi}(e)$ leads easily to

$$(\text{deltahedron}) \quad (\Gamma_T + \Gamma_e) \times \Gamma_{\sigma}(e) = (\Gamma_T + \Gamma_R) \times \Gamma_{\sigma}(f) \quad (7)$$

For every operation except the identity, both sides of (7) vanish, and under the identity both sides have character $4e = 6f = 12v - 24$; they are thus multiples of the regular representation

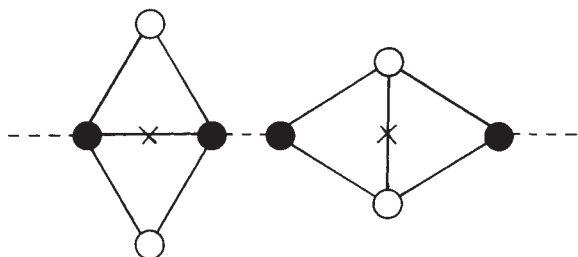


FIG. 2 Intersection of a reflection plane (dashed line) with the surface of a deltahedron. Linked units in the chain may be perpendicular or parallel kites, each contributing +1 to $\chi_{\sigma}(e)$ and $\chi_{\sigma}(v)$. Symbols as in Fig. 1.

of the molecular point group. For the platonic deltahedra (tetrahedron, octahedron and icosahedron) each side of (7) is equal to the regular representation itself.

Because of the dual relationship between deltahedra and 3-connected polyhedra⁴ (the class that includes the fullerene skeletons) the corresponding theorems are obtained by interchanging $\Gamma_{\sigma}(f)$ and $\Gamma_{\sigma}(v)$, $\Gamma_{\perp}(e)$ and $\Gamma_{\parallel}(e)$. Thus

$$(3\text{-connected}) \quad \Gamma_{\sigma}(f) \times \Gamma_T - \Gamma_T - \Gamma_R = \Gamma_{\sigma}(e) \quad (4')$$

$$(3\text{-connected}) \quad \Gamma_{\sigma}(v) \times \Gamma_T = \Gamma_{\parallel}(e) + \Gamma_{\sigma}(e) \quad (5')$$

$$(3\text{-connected}) \quad \Gamma_{\sigma}(v) \times \Gamma_R = \Gamma_{\perp}(e) + \Gamma_{\sigma}(e) \times \Gamma_e \quad (6')$$

$$(3\text{-connected}) \quad (\Gamma_T + \Gamma_e) \times \Gamma_{\sigma}(e) = (\Gamma_T + \Gamma_R) \times \Gamma_{\sigma}(v) \quad (7')$$

These results provide the group-theoretical basis underlying the classification made by Mingos and Wales¹³ of the skeletal orbitals in 3-connected cages. The atomic orbitals consist of one σ and two π orbitals at each vertex and transform as $\Gamma_{\sigma}(v) \times \Gamma_T$. From (5'), this basis splits into edge-bonding and edge-antibonding halves, transforming respectively as $\Gamma_{\sigma}(e)$ and $\Gamma_{\parallel}(e)$. Both parts can be further analysed, using (2) and (4'). For the bonding half this yields

$$(3\text{-connected}) \quad \Gamma_{\sigma}(e) = \Gamma_{\sigma}(f) + [\Gamma_{\pi}(f) - \Gamma_T - \Gamma_R] \quad (8)$$

The orbitals transforming as $\Gamma_{\sigma}(f)$ are bonding along edges and faces and thus are strongly stabilized. The remainder of the $\Gamma_{\sigma}(e)$ set has π symmetry at the face centres and is thus face-antibonding, resulting in overall non-bonding character. Thus the n -vertex cage is 'electron-precise', with $n/2 + 2$ strongly bonding and $n - 2$ weakly or non-bonding valence orbitals. Applying (2) to the edge-antibonding half yields:

$$(3\text{-connected}) \quad \Gamma_{\parallel}(e) = [\Gamma_{\sigma}(f) - \Gamma_0] \times \Gamma_e + [\Gamma_{\sigma}(v) - \Gamma_0] \quad (9)$$

The first bracket of (9) refers to edge-antibonds that form a rotation around the face centres, typical of a strongly antibonding interaction. The second bracket corresponds to the $L_{\sigma}(L > 0)$ orbitals of the TSH treatment⁷.

Euler's theorem for polyhedra is the dimensional part of more general character theorems that relate the symmetries of vertices, faces and edges. Their chemical significance lies in the way the theorems re-allocate properties of vertex atoms to edges or face centres. It should be noted that the tensorial nature of a property may change on such re-allocation; for example, polar vectors along edges may be combined to yield scalars at vertices and axial vectors on faces. This opens up new perspectives for the application of TSH theory. The physical background of the theorems is the continuity equation for fluid flow on a polyhedral surface. This equation may be of importance for the description of dynamic processes in cluster molecules. The theorems can also be read as purely group-theoretical lemmas pointing to the connection between representational orbits of point symmetry groups. Finally, we note that similar theorems can be developed for polyhedra that have toroidal topology. $\square$

Received 2 July; accepted 29 July 1991.

- Porterfield, W. W. *Inorganic Chemistry: A Unified Approach* (Addison-Wesley, Reading, Massachusetts, 1984).
- Kroto, H. W. *Nature* **329**, 529-531 (1987).
- Fowler, P. W. *Chem. Phys. Lett.* **131**, 444-449 (1986).
- Coxeter, H. S. M. *Convex Polytopes* 3rd edn (Dover, New York, 1973).
- Federico, P. J. *Descartes on Polyhedra* (Springer, New York, 1982).
- Fowler, P. W. & Quinn, C. M. *Theor. Chim. Acta* **70**, 333-350 (1986).
- Stone, A. J. *Inorg. Chem.* **20**, 563-571 (1981).
- Ceulemans, A. *Molec. Phys.* **54**, 161-181 (1985).
- Quinn, C. M., McKiernan, J. G. & Redmond, D. B. *J. Chem. Ed.* **61**, 572-579 (1984).
- Boyle, L. L. & Parker, Y. M. *Molec. Phys.* **39**, 95-109 (1980).
- Watson, J. K. G. *J. molec. Spectrosc.* **41**, 229-230 (1972).
- Marlin, W. (ed.) *The Artifacts of R. Buckminster Fuller—A Comprehensive Collection Of His Designs And Drawings* (Garland, New York, 1984).
- Mingos, D. M. P. & Wales, D. J. *Introduction to Cluster Chemistry* (Prentice-Hall, Englewood Cliffs, New Jersey, 1990).

ACKNOWLEDGEMENTS. A. C. acknowledges financial support from the Belgian National Science Foundation (NFWO) and the Belgian Government (Programmatie van het Wetenschapsbeleid).

Pressure-induced amorphization and disordering on cooling in a crystalline polymer

Sanjay Rastogi*, Mark Newman & Andrew Keller

H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK

* Permanent address: Department of Physics, University of Lucknow, Lucknow-226007, India

IN the course of exploring the phase diagram of the polymer poly(4-methyl-pentene-1) as a function of temperature and pressure by *in situ* X-ray diffraction, we have discovered some unusual phase behaviour. The polymer, crystalline under ambient conditions, becomes amorphous reversibly on increasing pressure in two widely separated temperature regimes ($\sim 20^\circ\text{C}$ and $\sim 200^\circ\text{C}$), the transformation occurring via a liquid-crystalline state. This suggests the possibility of 're-entrant' liquid-crystal and amorphous phases as pressure or temperature are varied. In the higher-temperature regime, the melting point shows a maximum as a function of pressure. The lower-temperature amorphous phase becomes crystalline on heating, and reverts to the glassy, disordered phase on cooling. Whereas pressure-induced amorphization has been observed previously in other systems, such as ice, silica and AlPO_4 , we know of no precedent for the disordering on cooling that we observe.

Crystals of the polyolefin P4MP1 normally contain helical chains with seven monomers for every two helical turns (7_2 helix). These helices form a tetragonal lattice with four chains per unit cell^{1,2}. The unit-cell parameters are $a = b = 18.65 \text{ \AA}$, c (chain direction) $= 13.76 \text{ \AA}$ and $\alpha = \beta = \gamma = 90^\circ$. The rather unusual fourfold coordination of helices has been attributed to the intermeshing requirement of the side groups², resulting in rather open packing. The calculated (tetragonal) crystal density (ρ_C) at room temperature is $\rho_C = 0.813$; this is not only low for a polyolefin but is also lower than the density of the corresponding amorphous state (ρ_A), which, extrapolated to room temperature (amorphous P4MP1 is not obtainable at this temperature), has the value $\rho_A = 0.830$. The significance of this similarity in densities, and in particular the inverted relationship between ρ_C and ρ_A , will become clear later. The melting point (T_m) of P4MP1 is 245°C at atmospheric pressure. There is thought to be a glass transition at $T_g = 30\text{--}50^\circ\text{C}$.

We examine here the effect of pressure and temperature on

the phase behaviour of P4MP1 by *in situ* X-ray diffraction, using high-pressure apparatus developed by Hikosaka³. The maximum pressure presently attainable is 6 kbar, and the temperature could be varied between room temperature and 300°C . For the X-ray diffraction studies a rotating-anode generator was used with a Mo target. The diffraction patterns were recorded as flat-plate photographs. They also contain reflections from the diamond windows of the pressure cell, which provide convenient internal standards for both spacing and intensity. The P4MP1 samples were mostly in the form of oriented films, the diffraction patterns from which have more information to convey than those of randomly oriented samples. Changes in molecular orientation as a function of pressure and temperature provide useful additional information.

We present results for selected pathways through the pressure-temperature phase diagram which involve transformations between the tetragonal solid (crystal) phase (C_t), the liquid (amorphous) phase (A) and liquid-crystalline phases (LC), these transformations being fully reversible. These pathways are indicated in the P - T phase diagram of Fig. 1, in which the phase boundaries are deduced from diffraction patterns such as those in Figs 2-4. The intervention of a second solid (crystal) phase, largely irreversibly, will be described elsewhere.

Pathway 1: increasing P at 20°C . As pressure increases, the diffraction peaks broaden and become diffuse, those close together merging while others disappear. Clearly, long-range crystal periodicity is lost and the system becomes disordered. The sequence is fully reversible with P (compare Fig. 2a and e). At the highest accessible pressure (~ 6 kbar), the sample is essentially an oriented amorphous material. In such oriented samples, retention of a certain amount of order in the packing of chains becomes apparent from the greater distinctness and intensity of the diffraction features along the equator (perpendicular to the orientation direction). At intermediate pressures the sample has the characteristics of an oriented, nematic liquid-crystal (LC) phase. Thus we have a transition sequence $C_t \leftrightarrow LC \leftrightarrow A$. The sequence of disordering (or ordering on pressure removal) is continuous—beyond a threshold of ~ 2 kbar there is no abrupt change from one distinct state to another. This prevents us from being able to define clear phase boundaries. This change from order to disorder with increasing pressure is consistent with the anomalous density relation $\rho_C < \rho_A$ referred to above. In this respect, P4MP1 seems to fall in the same class as water and several other materials, for which pressure causes a decrease (as opposed to the usual increase) in the melting point, but here the change occurs at a temperature $\sim 220^\circ\text{C}$ below the melting point at atmospheric pressure.

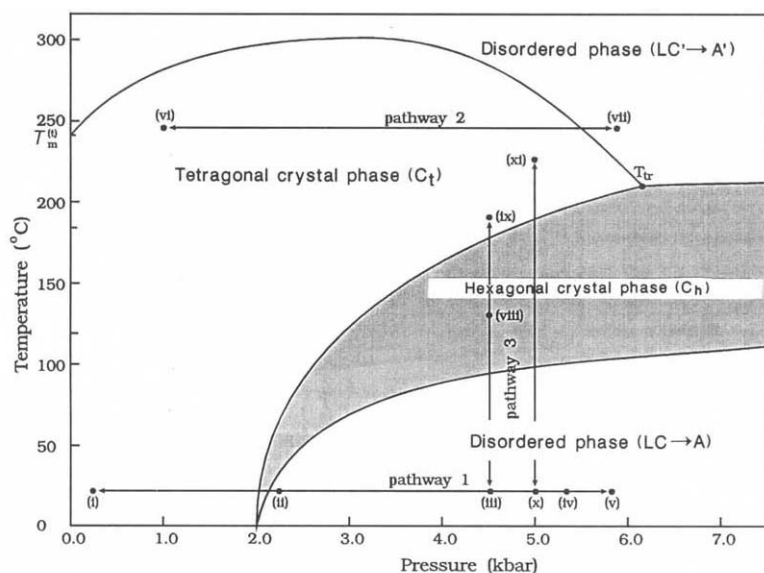
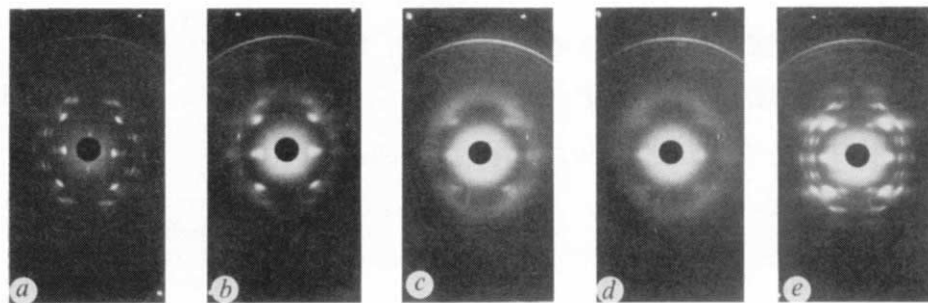


FIG. 1 P - T phase diagram of poly(4-methyl-pentene-1) (P4MP1), showing the phases between which fully reversible transformations occur: liquid (amorphous), solid (crystal in tetragonal form) and intermediate (liquid crystal) denoted by A, C_t and LC respectively. Roman numerals denote experimental points documented by the X-ray diffraction patterns used here as examples (see text). There is continuous gradation between LC and A, and we therefore represent the two collectively as LC-A, corresponding to disordered structures, as compared with C_t . The reversible transition from LC-A to C_t on raising or lowering T involves passing through a broad regime (shaded) which encompasses a further (hexagonal) crystal phase, C_h . This phase may be encountered on heating along pathway 3 (iii-viii). A transition from C_h to C_t occurs on further heating (viii-ix). In contrast, on cooling at fixed P , C_t goes directly in the disordered phase without the intervention of the C_h phase (ix-iii). The existence of a triple point is implied by the phase diagram.

FIG. 2 X-ray diffraction patterns of drawn films of P4MP1 at room temperature recorded *in situ* at pressures of (pathway 1 in Fig. 1) *a*, 0.24 kbar, *b*, 2.24 kbar, *c*, 4.56 kbar, *d*, 5.36 kbar and *e*, 0.24 kbar. *a* to *d* correspond to points (i) to (iv) respectively in Fig. 1; *e* corresponds to a reversal to point (i). The sharp reflections on the periphery here and subsequently in Figs 3 and 4 are due to diamond; the rings arise from a gasket. These reflections serve as intensity standards when comparing different members in a sequence.



Pathway 2: increasing P at constant $T = 248^\circ\text{C}$. (Note that the melting point T_m as a function of P has been examined calorimetrically up to 2 kbar in refs 4, 5.) At this temperature, the disordered state could be reached at sufficiently high P . Figure 3*a* corresponds to the crystalline and *b* to the disordered state, with disordering setting in continuously beyond a threshold P corresponding to the point at which pathway 2 crosses the phase boundary. The process is reversible, the C_i state being regained on reducing P (Fig. 3*c*). Here, in contrast to pathway 1, passage through the LC' or A' region is accompanied by loss of orientation of the chains.

It follows from the results (and similar ones not described here) that at sufficiently high P , $dT_m/dP < 0$. Accordingly, the plot of T_m against P should have a maximum at some value $P \equiv P'$. Thus we have a situation where for $P < P'$, T_m rises with P , as for most substances, reaches a maximum at P' and decreases thereafter.

Pathway 3: T lowered isobarically from high to room temperature at high P . In this case the sample, highly crystalline at the elevated temperature, loses order; that is, it changes towards the liquid crystalline or amorphous states on cooling. The extent of the disordering depends on the pressure: the same decrease in T leads to a more pronounced disorganization for higher P . Figure 4 shows the effect on cooling for two different pressures. The disordered end states follow a sequence similar to that in Fig. 2*a-d*. The process can again be regarded as reversible, insofar as the crystalline state (C_i) is regained on heating at constant pressure; but this thermal cycle cannot be carried out without the intervention of a further, different crystal phase (to be discussed elsewhere).

The most salient features of these results are the transformation of a crystalline polymer phase into a disordered one on application of pressure and the observation of ordering on

heating and disordering on cooling. To our knowledge, the latter has not been reported before for any material (except for a small effect in carbon⁶). Disorder with increase in pressure is unusual but not unprecedented—similar behaviour has been recorded for AlPO_4 (refs 7–9), SiO_2 (quartz)¹⁰, SnI_4 , LiKSO_4 , $\text{Ca}(\text{OH})_2$ (refs 11–14), ice^{15–17}, $\text{CaAl}_2\text{Si}_2\text{O}_8$ (anorthite) and Fe_2SiO_4 (fayalite)¹⁸. Although most of these transformations are irreversible, in some cases reversibility has been observed (SnI_4 , AlPO_4 , LiKSO_4 , $\text{Ca}(\text{OH})_2$). For AlPO_4 the initial crystal orientation was also regained on reversing the transformation ('memory glass')⁹.

Isothermal pressure-induced amorphization of SiO_2 quartz¹⁰ was explained on the basis of the higher compressibility of the melt or amorphous state compared with that of the crystal: at a sufficiently high pressure, this makes the amorphous state denser than the crystalline state. T_m should rise with increasing P while $\rho_C > \rho_A$, but will decrease again when $\rho_C < \rho_A$. It is suggested that amorphization sets in as P is increased iso-

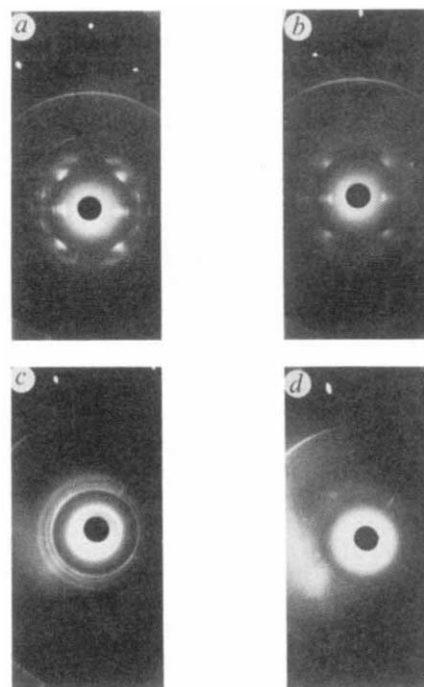


FIG. 4 X-ray diffraction patterns of P4MP1 films showing the effect of isobaric cooling to room temperature at two different pressures (pathway 3 in Fig. 1). *a*, 4.4 kbar, at a starting temperature of 200°C ; drawn film. *b*, as *a*, after cooling to room temperature. *a* and *b* correspond to points (ix) and (iii) in Fig. 1. *c*, 5 kbar at a starting temperature of 220°C ; unoriented film. *d*, as *c* after cooling to room temperature. *c* and *d* correspond to points (xi) and (x) in Fig. 1.

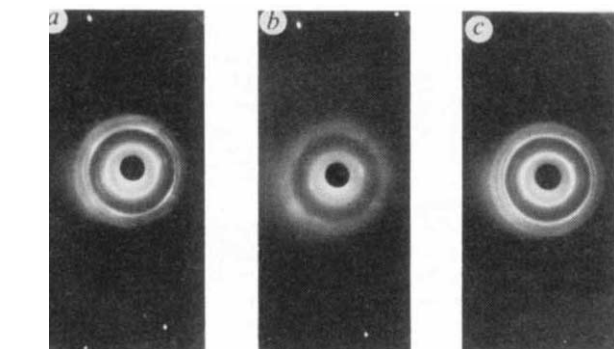


FIG. 3 X-ray diffraction patterns of drawn films at higher temperatures (partially disordered on heating) of P4MP1 at $\sim 248^\circ\text{C}$ at pressures of *a*, 1 kbar; *b*, 6 kbar; *c*, ~ 0.74 kbar (pathway 2 in Fig. 1). *a* and *b* correspond to points (vi) and (vii) respectively in Fig. 1, and *c* is a reversal to point (vi).

thermally when it intersects the downward sloping branch of the T_m against P curve at $P > P'$. The same has been proposed for the water-ice system¹⁵⁻¹⁷, in which hexagonal ice is transformed into a high-density amorphous phase at -196°C and $P \approx 10$ kbar. The explanation is consistent with amorphization (melting) of P4MP1 at temperatures of 220 – 250°C , where $\rho_C > \rho_A$ at atmospheric pressure; but in contrast to SiO_2 and ice, the process is reversible.

The occurrence of pressure-induced amorphization of P4MP1 at 20°C is consistent with the fact that $\rho_C < \rho_A$ at this lower temperature (ρ_C and ρ_A are known to cross over with temperature). Disorder sets in at the comparatively low pressure of 2 kbar. Given the stability of the crystal phase at higher temperature at this and higher pressures, this means that the disordered phase is re-entrant.

The exact nature of the disordered phase is uncertain. As the crystal phase always intervenes between the two disordered regions (at $T > 200^\circ\text{C}$ and $T < 20^\circ\text{C}$), it is difficult to ascertain whether they are identical; but they should be different at least to the extent that a glass is expected at low temperatures but a liquid at high T . Further, the question arises as to whether or not the low- T disordered phase is an equilibrium state. The reversibility suggests that it is, but retention of orientation in the crystal formed on heating indicates that full randomization may not occur during the preceding cooling. Preservation of some orientation is a common feature in polymer systems, owing to slow chain relaxation in the highly viscous amorphous state and thus there may be no need to invoke a specific nonequilibrium situation, as is the case for AlPO_4 (ref. 9). Even so, we recognize that dynamic instabilities^{9,18} or frustrated solid-state transitions cannot be excluded as possible sources of disorder.

We note that an equilibrium phase transition from a crystal to a disordered state on cooling does not violate thermodynamics despite the decrease in entropy associated with such a process. Systems exist in which entropy increases on ordering, such as 'crystallization' of colloidal latex spheres¹⁹. For polymers with loosely packed crystal structures, as compared with their denser, configurationally hindered melts (or glasses), a similar situation is conceivable. In fact, this behaviour was predicted by Tamman^{20,21}, who envisaged maxima of phase boundaries, as a function of T and P , for single-component systems. He discussed phase boundaries comprising closed loops enclosing a crystal phase, surrounded by an amorphous region (liquid or glass). According to Tamman (and Vogel²²), these predictions are likely to require extreme conditions for their realization, not available in a terrestrial laboratory. Yet, as it appears, they may have become realizable closer to standard laboratory conditions for the polymer in our experiments. $\square$

Stable carbon isotope variation in C_3 and C_4 plants along the Amazon River

Luiz A. Martinelli^{*†}, Allan H. Devol^{*},
Reynaldo L. Victoria[†] & Jeffrey E. Richey^{*}

^{*} School of Oceanography, WB-10, University of Washington, Seattle, Washington 98195, USA

[†] Escola Superior de Agricultura 'Luiz de Queiroz', Piracicaba SP 13400, Brazil

ALL plants assimilate ^{12}C in preference to ^{13}C . As a result of this isotope fractionation, the tissues of subaerial plants have lower $^{13}\text{C}/^{12}\text{C}$ ratios than that of atmospheric CO_2 . By contrast, plant respiration and tissue decomposition are accompanied by little, if any, fractionation and hence release ^{13}C -depleted biogenic CO_2 back into the atmosphere. If this biogenic CO_2 is reassimilated before it is thoroughly mixed into the atmosphere, a further depletion of ^{13}C in the plant tissue will result. This recycling effect has been found most often in vertical variations of stable isotope composition in tropical forests where plant tissues near the forest floor are more depleted in ^{13}C (refs 1–5). Here we show that the intensity of biogenic CO_2 recycling in flood plain forests of the Amazon systematically increases inland, in the western Amazon basin. We also show that a similar recycling mechanism affects the ^{13}C composition of semiaquatic grasses owing to evasion of biogenic CO_2 from the Amazon river. But in this case the degree of recycling is more pronounced in the eastern basin because the flux of $^{13}\text{CO}_2$ out of the river is smaller there. Our data indicate that significant spatial carbon isotope gradients can exist across the same general ecosystem, both between different species and also within a single species. Recycling effects therefore need to be taken into account in studies that try to relate plant carbon composition to animal and human diet, and in those attempting to determine the carbon isotope composition of the ancient atmosphere from preserved plant tissues.

Because of different pathways in photosynthetic biochemistry, C_3 and C_4 plants discriminate against the heavy carbon isotope (^{13}C) (refs 6, 7). Consequently, tissues of C_3 plants such as trees have an average $\delta^{13}\text{C}$ value of -27% whereas C_4 plants, which are mainly grasses, average about -12% (ref. 8). Within both plant categories, genetic and environmental factors^{9,10} cause more subtle variability in the carbon isotopic composition. The isotopic composition of the CO_2 source is one of the most important environmental factors. Here we show that there are clear trends in the $\delta^{13}\text{C}$ of C_3 and C_4 plants of the Amazon River flood plain (*várzea*), which result from the systematic variation in the biogenic CO_2 contribution to local atmospheric CO_2 composition on a regional scale (1,800 km).

Plant samples from the Amazon River *várzea* were collected along an 1,800-km west-to-east (upstream to downstream) transect between Vargem Grande ($3^\circ 16' \text{S}$, $67^\circ 55' \text{W}$) and Obidos ($1^\circ 55' \text{S}$, $55^\circ 30' \text{W}$), Brazil (Fig. 1). C_3 plants were collected during November and December 1988 from different *várzea* sites. At each site, we sampled leaves from at least five different tree species and five different understory plant species. All tree leaves were collected in the forest 100–400 m from the river at a height of ~ 10 m, and the understory leaves were collected at a height less than 1 m. C_4 plant samples were restricted to *Echinochloa polystachya*, a tropical semiaquatic grass with emergent leaves¹¹. Grass leaf samples from 2–3 plants growing in flooded areas were collected as the water began to rise (October–November 1983) and as it began to fall (June–July

Received 15 April; accepted 26 July 1991.

- Natta, G., Corradini, P. & Bassi, I. W. *R.C. Acad. Lincei* **19**, 404–411 (1955).
- Frank, F. C., Keller, A. & O'Connor, A. *Phil. Mag.* **4**, 200–214 (1959).
- Hikosaka, M. & Seto, T. *Jap. J. appl. Phys.* **21**, L332–L334 (1982).
- Jain, P. C., Wunderlich, B. & Chaubey, D. R. *J. Polym. Sci. Polym. Phys.* **15**, 2271–2275 (1977).
- Xoller, P., Starkweather, H. W. Jr & Jones, G. A. *J. Polym. Sci. Polym. Phys.* **24**, 1451–1458 (1986).
- Zemansky, M. W. *Heat and Thermodynamics* 5th Edn, 376 (McGraw-Hill, New York, 1968).
- Jayaraman, A., Wood, D. L. & Maines, R. G. *Phys. Rev.* **B35**, 8316–8321 (1987).
- Williams, Q. & Jeanloz, R. *Science* **239**, 902–905 (1988).
- Kruger, M. B. & Jeanloz, R. *Science* **249**, 647–649 (1990).
- Hemley, R. J., Jephcoat, A. P., Mao, H. K., Ming, L. C. & Manghnani, M. H. *Nature* **334**, 52–54 (1988).
- Fujii, Y., Kowaka, M. & Onodera, A. *J. Phys.* **C18**, 789–797 (1985).
- Sugai, S. *J. Phys.* **C18**, 799–808 (1985).
- Mei, F. E. A., Lemos, V. E., Cerdeira, F. & Mendes, F. J. *Phys. Rev.* **B35**, 3633–3636 (1987).
- Sankaran, H., Sikka, S. K., Sharma, S. M. & Chidambaram, R. *J. Phys. Rev.* **B38**, 170–173 (1988).
- Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **310**, 393–395 (1984).
- Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **314**, 76–78 (1985).
- Hemley, R. J., Chen, L. C. & Mao, H. K. *Nature* **338**, 638–640 (1989).
- Williams, Q. & Jeanloz, R. *Nature* **338**, 413–415 (1989).
- Baus, M. J. *J. stat. Phys.* **48**, 1129–1146 (1987).
- Tamman, G. *Kristallisieren und Schmelzen* (Johann Ambrosius Barth, Leipzig, 1903).
- Tamman, G. & Mehl, R. F. *The States of Aggregation* (Van Nostrand, New York, 1925).
- Vogel, R. *Die Heterogene Gleichgewichte* 2nd Edn (Academische, Leipzig, 1959).

ACKNOWLEDGEMENTS. We thank M. Hikosaka for the experimental apparatus, and R. Koningsveld, R. Evans, A. Toda and Sir Charles Frank for advice and discussions. Support by the SERC is gratefully acknowledged.

[†] Permanent address: Centro de Energia Nuclear na Agricultura, Av. Centenario 303, Piracicaba SP 13400, Brazil.

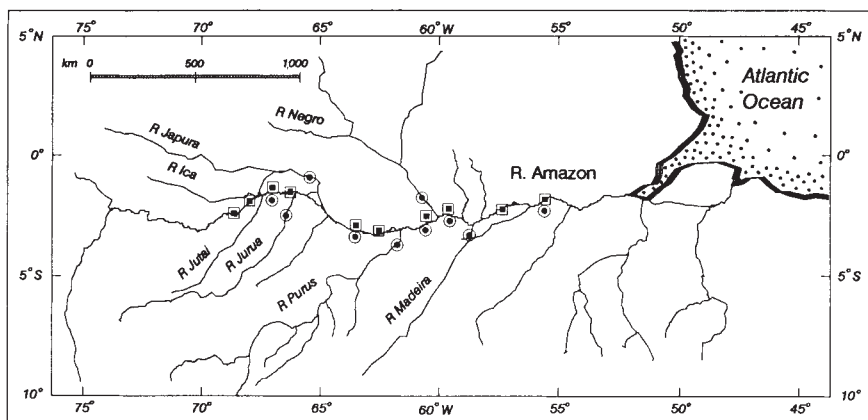


FIG. 1 Study area. Sampling points are indicated by squares (C_4 grasses) and by circles (C_3 plants).

1984). As not all parameters or samples are independent, we have used nonparametric statistical analyses (the Spearman Rank test for differences within groups, and the Wilcoxon test for differences between groups). All confidence intervals are at the 95% level.

The $\delta^{13}\text{C}$ of C_3 plant tissues increased downriver, whereas the $\delta^{13}\text{C}$ of C_4 grasses decreased downriver (Fig. 2). The overall increase in $\delta^{13}\text{C}$ between the sites furthest upstream (Vargem Grande) and downstream (Obidos) was $\sim 3\%$ for tree leaves and $\sim 4\%$ for understory leaves. For grass leaves, the corresponding decrease in $\delta^{13}\text{C}$ was 1.5% .

One possible cause for the observed downriver increase of

$\delta^{13}\text{C}$ of C_3 tissues is a parallel variation in the isotopic composition of atmospheric CO_2 . Indeed, Longinelli and Edmond¹² observed a downriver increase in the $\delta^{13}\text{C}$ of atmospheric CO_2 ($\delta^{13}\text{CO}_2\text{-atm}$) sampled over the Amazon of almost 10‰ during the dry season in 1976 between Iquitos, Peru and the Atlantic. They attributed this difference to an increase downriver in biogenic CO_2 produced by plant respiration and by decay of organic matter. They did not, however, observe such a gradient during the 1977 rainy season. They also caution that their sampling technique for atmospheric CO_2 may have introduced considerable isotope fractionation. The data of Quay *et al.*¹³ do not indicate any consistent difference in the $\delta^{13}\text{CO}_2\text{-atm}$ collected above the river at locations between Vargem Grande and Obidos (Fig. 3). Thus, an inland decrease in $\delta^{13}\text{CO}_2\text{-atm}$ may well not exist.

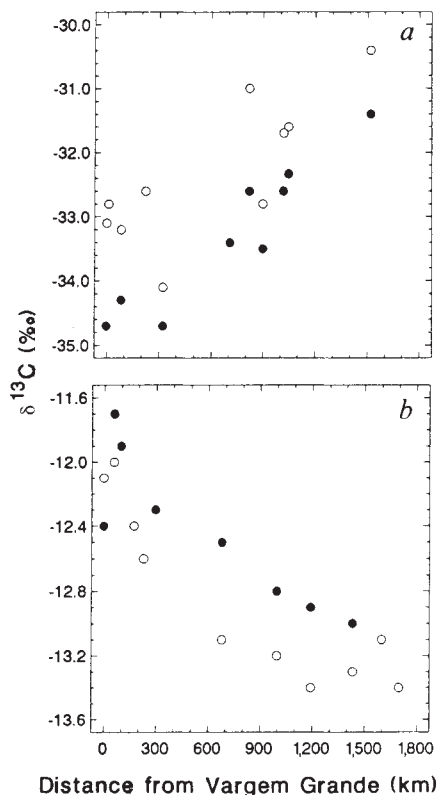


FIG. 2 Downriver variation of the carbon isotopic composition of *a*, C_3 plant leaf samples: $\circ$, tree leaves; $\bullet$, understory leaves. *b*, C_4 plant leaf samples, $\circ$ taken during falling water; $\bullet$, rising water. The samples were air-dried and ground for further analyses. The carbon isotope composition was determined by the combustion of the vegetal material in evacuated vials and analysis of the evolved CO_2 in a Micromass 602 E mass spectrometer. The results are expressed as $\delta^{13}\text{C} = [({}^{13}\text{C}/{}^{12}\text{C})_{\text{sample}} / ({}^{13}\text{C}/{}^{12}\text{C})_{\text{PDB}} - 1] \times 1,000$.

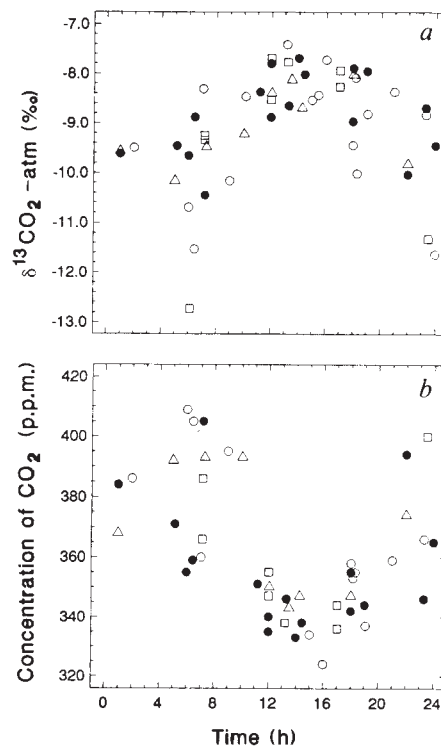


FIG. 3 *a*, Variation of $\delta^{13}\text{C}$ of the atmospheric CO_2 ($\delta^{13}\text{CO}_2\text{-atm}$) and *b*, atmospheric CO_2 concentration throughout the day collected over the river (3 m height) in four different regions: Vargem Grande, 0 km in Fig. 2 ($\bullet$); Itapeua (Δ), 687 km; Manacapuru, 992 km ($\circ$) and Obidos, 1,696 km ($\square$). The graph is constructed from data for different cruises presented in ref. 13. Note that there is a systematic daily variation although there is no apparent downstream trend.

An alternative explanation for the observed gradient in C_3 plant tissue is that inland plants are exposed to the ^{13}C -depleted biogenic CO_2 that is produced during the night (Fig. 3) for more of the morning photoperiod than those near the ocean. The depletion of ^{13}C seen in Fig. 3 results from the trapping of biogenic CO_2 produced during night respiration in the nocturnal, tropospheric boundary layer formed through near-surface cooling of the atmosphere¹⁴. The return of $\delta^{13}\text{C}$ and CO_2 concentration to values characteristic of the bulk atmosphere takes place in the morning, when solar heating breaks up the boundary layer through thermal convection¹⁴. The break up of the nocturnal boundary layer occurs, on average, later in the day in the western (upstream) part of the Amazon basin than in the east¹⁴. Amazon trees take up proportionally more CO_2 during the morning hours than during the afternoon¹⁵. These two observations suggest that plants growing in the western portion of the basin would fix a higher proportion of nocturnally produced biogenic CO_2 , which would in turn result in their tissues being more depleted in ^{13}C than those of plants in the east.

The change of the $\delta^{13}\text{C}$ values of C_4 grasses downriver was in the opposite direction to that of C_3 plants. *E. polystachya* is a semiaquatic macrophyte, and its leaves are therefore close to the river surface¹¹. The Amazon river is supersaturated with

CO_2 by about a factor of 10 (refs 16, 17) with downriver water being 1.5–2.0 times more supersaturated than upstream. Consequently, the flux of CO_2 from the river to the atmosphere must also increase downstream (Fig. 4). The CO_2 supersaturation of the river results from respiratory input of biogenic CO_2 , which is depleted in ^{13}C relative to the atmosphere. Immediately above the river surface, therefore, the dilution of atmospheric CO_2 by light CO_2 evading the river is greater downriver than upriver. This is probably the main cause of the decrease in $\delta^{13}\text{C}$ of C_4 grasses downriver.

The $\delta^{13}\text{C}$ of the CO_2 evading the river ($\delta^{13}\text{C}_{\text{CO}_2\text{-aq}}$) also decreases downstream. Using the data of Quay *et al.*¹⁸, we calculated the isotopic compositions of this CO_2 (Fig. 4). The $\delta^{13}\text{C}_{\text{CO}_2\text{-aq}}$ values decrease downriver by $\sim 1.5\%$ during both the rising and the falling water cruises (Fig. 4). Consequently, aquatic grasses growing further downstream are bathed by increasingly greater fluxes of CO_2 , more depleted in ^{13}C , from the river.

Our results show that gradients in carbon isotopic compositions expressed by vegetation over long distances can be generated by incremental change in physical processes affecting the extent of assimilation of locally derived biogenic CO_2 . Such gradients have been previously observed only in isolated forest stands, but not in aquatic environments. At an extreme upstream location, van der Merwe and Medina¹⁹ observed very negative $\delta^{13}\text{C}$ values in higher organisms of a rain forest foodweb in the Orinoco river basin (northwest of Amazon). They attributed these to the depleted $\delta^{13}\text{C}$ of forest plants, caused in turn by the biogenic CO_2 recycling mechanism discussed earlier. This depletion is of the same magnitude as our observed downstream gradient in C_3 plant tissues. Consequently, in the upper Amazon one might expect to find relatively negative $\delta^{13}\text{C}$ values for higher organisms, as was observed by van der Merwe and Medina¹⁹, but heavier (less negative) values downstream, in the lower Amazon.

Recently, Marino and McElroy²⁰ reported a variation in $\delta^{13}\text{C}$ of $\sim 1.5\%$ for nitrate cellulose from kernels of maize (a C_4 grass) cultivated between 1948 and 1986 in Iowa, USA. As the maize was all grown in the same fields, the authors attributed the variation to changes in the partial pressure of CO_2 in the atmosphere. This explanation is probably correct for their data. Our data for *E. polystachya* show a similar variation of $\sim 1.5\%$ and imply that the variability caused by assimilation of locally derived biogenic CO_2 could produce an effect similar to that resulting from overall changes in the global atmospheric partial pressure of CO_2 . Caution should therefore be used when comparing and interpreting carbon isotope compositions to deduce palaeoatmospheric conditions, even when the same plant species are involved. □

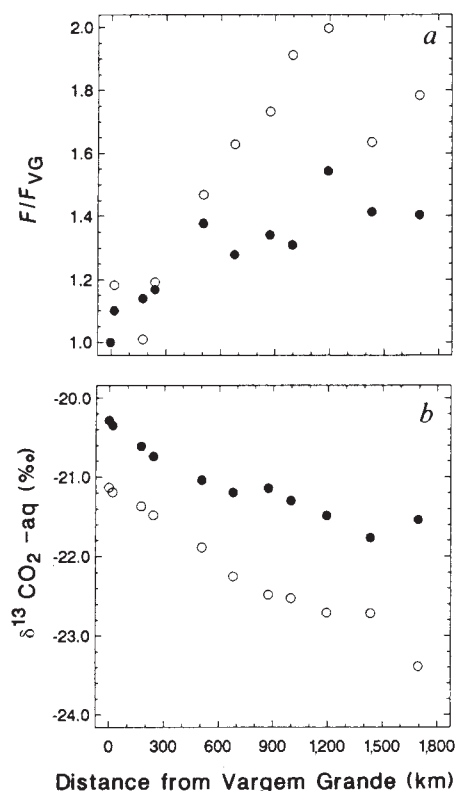


FIG. 4 *a*, Downriver variation of the ratio of the CO_2 flux from the river to the atmosphere (F) to that at Vargem Grande (F_{VG}). The ratio (F/F_{VG}) is therefore 1.0 at Vargem Grande (0 km). We calculated the fluxes assuming a constant gas exchange coefficient¹⁶. *b*, Downriver variation of $\delta^{13}\text{C}$ for CO_2 evolved from the Amazon River ($\delta^{13}\text{C}_{\text{CO}_2\text{-aq}}$) calculated by assuming that measured alkalinity was due only to carbonates and by using the following mass and isotope balance equations: $\text{DIC} = \text{CO}_2 + \text{Alk}$ (equation 1); $\text{DIC} \times \delta^{13}\text{C}_{\text{DIC}} = (\text{CO}_2\text{-aq} \times \delta^{13}\text{C}_{\text{CO}_2\text{-aq}}) + (\text{Alk} \times \delta^{13}\text{C}_{\text{Alk}})$ (equation 2) where DIC, total dissolved inorganic carbon; $\text{CO}_2\text{-aq}$, the dissolved CO_2 concentration; Alk, alkalinity ($\text{HCO}_3^- + 2\text{CO}_3^{2-}$) in the river water. Within the pH range of the Amazon River in our study area (pH 6–7) alkalinity is almost entirely due to HCO_3^- . Thus, by assuming that the isotope fractionation between dissolved CO_2 and HCO_3^- is 8.7‰ at 27 °C, equation 2 can be solved for the $\delta^{13}\text{C}$ of the dissolved CO_2 ($\delta^{13}\text{C}_{\text{CO}_2\text{-aq}}$): $\delta^{13}\text{C}_{\text{CO}_2\text{-aq}} = [(\text{DIC} \times \delta^{13}\text{C}_{\text{DIC}}) - (\text{Alk} \times 8.7)] / \text{DIC}$ (equation 3). In both panels ○ are for falling water and ● are for rising water.

Received 16 April; accepted 16 July 1991.

- Vogel, J. C. *Ecol. Plant.* **13**, 9–94 (1978).
- Medina, E. & Minchin, P. *Oecologia (Berl.)* **45**, 377–378 (1980).
- Medina, E., Montes, G., Cuevas, E. & Rokzandic, Z. *J. trop. Ecol.* **2**, 207–217 (1986).
- van der Merwe, N. J. & Medina, E. *Geochim. cosmochim. Acta* **53**, 1091–1094 (1989).
- Sternberg, L. S. L., Mulkey, S. S. & Wright, S. J. *Ecology* **70**, 1317–1324 (1989).
- Emrich, K., Ehalt, D. H. & Vogel, J. C. *Earth planet. Sci. Lett.* **8**, 363–371 (1970).
- Mook, W. G., Bommerson, J. C. & Staverman, W. H. *Earth planet. Sci. Lett.* **22**, 169–176 (1974).
- Smith, B. N. & Epstein, S. *Plant Physiol.* **47**, 380–384 (1971).
- Vogel, J. C. in *Fractionation of the Carbon Isotopes During Photosynthesis* (Springer, New York, 1980).
- Farquhar, G. D., O'Leary, M. H. & Berry, J. A. *Aust. J. Plant Physiol.* **9**, 121–138 (1982).
- Junk, W. in *The Amazon Limnology and Landscape Ecology of a mighty Tropical River and its Basin* (ed. Sioli, H.) 271–293 (Junk, Dordrecht, 1984).
- Longinelli, A. & Edmond, J. M. *J. geophys. Res.* **88**, 3703–3717 (1983).
- Quay, P., King, S., Wilbur, D., Wofsy, S. & Richey, J. *J. geophys. Res.* **94**, 18327–18336 (1990).
- Wofsy, S. C., Harris, R. C. & Kaplan, W. A. *J. geophys. Res.* **93**, 1377–1387 (1988).
- Fan, S. M., Wofsy, S. C., Bakwin, P. S., Jacob, D. J. & Fitzjarrald, D. R. *J. geophys. Res.* **95**, 16851–16864 (1990).
- Devol, A. H., Richey, J. E., Quay, P. D. & Martinelli, L. *Limnol. Oceanogr.* **32**, 235–248 (1987).
- Richey, J. E. *et al. Limnol. Oceanogr.* **35**, 352–371 (1990).
- Quay, P. *et al. Limnol. Oceanogr.* (in the press).
- van der Merwe, N. J. & Medina, E. *J. archaeol. Sci.* **18**, 249–259 (1991).
- Marino, B. & McElroy, J. *Nature* **349**, 127–131 (1991).

ACKNOWLEDGEMENTS. We thank the crew of the RV L. M. Amanai and P. B. de Camargo and M. Piccolo for help in the field sampling. We thank J. Hedges, P. Quay, L. Sternberg and D. Hastings for comments.

Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle

Kenneth Mopper^{*†}, Xianliang Zhou^{*‡},
Robert J. Kieber^{*‡}, David J. Kieber^{*‡},
Richard J. Sikorski^{*} & Ronald D. Jones[†]

^{*} Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149, USA

[†] Drinking Water Research Center and Department of Biological Sciences, Florida International University, Miami, Florida 33199, USA

THE processes that regulate the cycling of oceanic dissolved organic carbon (DOC), one of the largest carbon reservoirs on the Earth's surface¹, are largely unknown. DOC residues in the deep sea, below 500 m, seem to be composed mainly of biologically refractory compounds^{2–10} such as humic substances¹¹. The average apparent ¹⁴C age of this refractory DOC is >6,000 yr in the deep Pacific², suggesting that its rate of turnover is slow, but the pathways and rates responsible for this apparent slow turnover are unknown. Several studies have shown that aquatic humic substances are photochemically degraded by sunlight into biologically labile and/or volatile organic compounds^{12–14} and carbon monoxide^{15,16}. Here we present new data which suggest that this photochemical degradation pathway is the rate-limiting step for the removal of a large fraction of oceanic DOC. This rate will increase with increasing flux of solar ultraviolet-B radiation. We estimate the oceanic residence time of biologically refractory, photochemically reactive DOC to be 500–2,100 yr, which is less than its average apparent ¹⁴C age. The injection of 'old carbon' from sediments into the deep sea may explain this discrepancy.

Several lines of evidence suggest that oceanic DOC can be categorized into different fractions by their ability to photochemical or chemical degradation procedures, being either more labile or more resistant^{2,7–10,17,18}. The more resistant fraction appears to be relatively more abundant in the upper water column and may be composed of recently produced biological macromolecules, colloids and sub-micrometre particles^{17,19}. The other DOC fraction seems to be biologically refractory^{3–10} and old, ranging in average apparent ¹⁴C age from 3,000–4,000 yr in the upper water column (based on extracted humic substances)¹⁰ to at least 6,000 yr in the deep sea^{2,10}, where it represents a large portion (40–80%) of the DOC pool^{10,18}. Although this old DOC resists biological degradation, it is readily attacked and degraded by oxidative photochemical or chemical procedures^{2,10,17,18} and by sunlight^{13–15}. Previously, we estimated an ocean residence time (τ) of 40,000 yr for this DOC based on attack of DOC by sunlight-produced hydroxyl radicals at the sea surface²¹. This τ is much older than the average apparent ¹⁴C age, which suggests that degradation by hydroxyl radicals alone does not play a significant part in ocean carbon cycling. Thus, we suggested that other, as yet unknown, photochemical pathways, possibly direct photolysis and photosensitized reactions²², are quantitatively much more important in degrading ocean DOC during its passage through the sunlit layer of the ocean.

Seawater samples were collected from the surface down to 4,000 m during five cruises between 1986 and 1989 to the south-western Sargasso Sea. Sampling, irradiation and analytical procedures are described in detail elsewhere^{14,15,23–26}. Briefly, irradi-

ation experiments for carbonyl and α -keto acid production rates were usually initiated within 2 h of sampling, and were done in triplicate using sealed quartz flasks with natural sunlight (solar noon, cloudless sky, 26° N) for periods of ≤ 5 h. During irradiation, the flasks were kept in a bath of circulating surface sea water for temperature control. Carbonyl compounds and α -keto acids were quantified by high-performance liquid chromatography^{23–25}. For most measurements of CO photochemical production, seawater samples were stored refrigerated in clear glass bottles and shipped back to Miami. Photochemical production rates of CO, carbonyl compounds and α -keto acids from rooftop laboratory irradiations agreed within 10% with ship-board irradiations of freshly collected samples. All results were corrected for dark controls, which were usually <10% of the photochemical production signal.

When seawater samples from different depths were exposed to sunlight, a variety of carbon compounds were produced (Table 1). Photochemical production rates were a few nM C h⁻¹, which is in agreement with rates from other ocean regions^{12,14,15,27}. Figure 1 shows typical results for formaldehyde; other photoproducts displayed similar profiles. The highest photochemical production rates were observed for samples from a depth of 800–1,000 m, corresponding to a maximum in DOM fluorescence (excitation 310–360 nm, emission 410–460 nm) (Fig. 1). Several studies have shown that humic substances are the dominant source for photochemically produced low-molecular-weight (LMW) carbon species in natural waters^{12,14–16}. In agreement with those studies, we found that photochemical production rates of LMW carbon species in Sargasso Sea samples were highly correlated with DOM fluorescence (for example, $r^2 = 0.91$, November cruise, Fig. 1b). DOM fluorescence in turn is strongly correlated with the humic concentration^{14,28,29}. Furthermore, the measured photochemical production rates in Sargasso Sea samples are in good agreement (within a factor of two) with rates predicted from irradiation experiments with sea water containing purified humic extracts¹⁴ (R.D.J., unpublished results). In those experiments, humic extracts from marine and terrestrial sources were added at various concentrations to Sargasso Sea water. The photochemical production rates of LMW carbonyl species and CO were found to be linearly related to the humic concentration and fluorescence. From the above evidence, we suggest that the photochemically produced carbonyl compounds, α -keto acids, and CO that were measured in most Sargasso Sea samples (20–4,000 m) were derived mainly from humic substances. For near-surface samples (0–20 m, Table 1), however, photochemical production rates were higher than that predicted from the DOM (humic) fluorescence of those waters. The reason for this result is not understood, but sunlight-induced release of

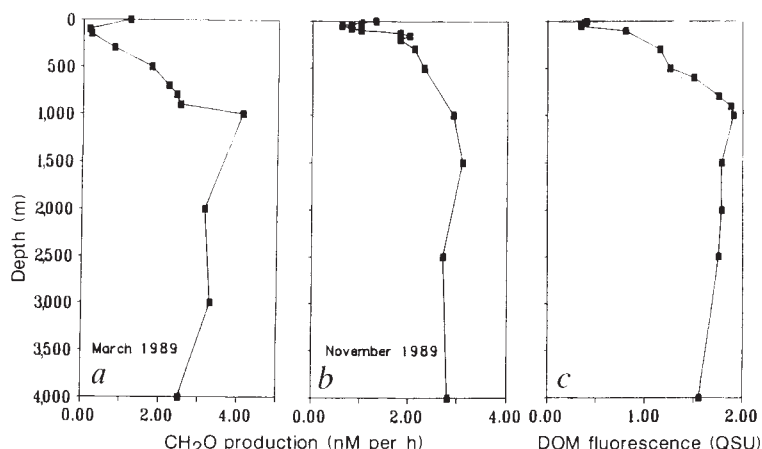
TABLE 1 Summary of photochemical production rates of LMW carbonyl compounds and CO in Sargasso Sea Waters measured on five cruises

Compound	Photochemical production rate (nM C per h)		
	1–20 m	20–150 m	500–4,000 m
Formaldehyde	1.0 $\pm$ 0.4 (n = 7)	0.5 $\pm$ 0.2 (n = 8)	2.4 $\pm$ 0.5 (n = 20)
Acetaldehyde	1.0 $\pm$ 0.5 (n = 4)	0.5 $\pm$ 0.3 (n = 6)	2.8 $\pm$ 0.8 (n = 15)
Glyoxal	0.7 $\pm$ 0.3 (n = 2)	0.4 $\pm$ 0.2 (n = 3)	1.1 $\pm$ 0.3 (n = 8)
Glyoxylate	0.8 $\pm$ 0.3 (n = 3)	0.8 $\pm$ 0.5 (n = 8)	2.3 $\pm$ 0.7 (n = 11)
Pyruvate	0.6 $\pm$ 0.2 (n = 3)	1.1 $\pm$ 0.4 (n = 8)	2.0 $\pm$ 0.5 (n = 11)
CO	18.6 $\pm$ 2.7 (n = 6)	9.6 $\pm$ 0.8 (n = 5)	16.1 $\pm$ 2.2 (n = 22)
Total	22.7 $\pm$ 2.8	12.9 $\pm$ 1.1	26.7 $\pm$ 2.6

n = number of samples; each sample was analysed in duplicate or triplicate; 1 σ standard deviation

[‡] Present addresses: Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA (K.M.); DAS/ECB, Brookhaven National Laboratory, Upton, New York 11973, USA (X.Z.); Department of Chemistry, University of North Carolina, Wilmington, North Carolina 28403, USA (R.J.K.); Department of Chemistry, College of Environmental Science and Forestry, State University of New York, Syracuse, New York 13210, USA (D.J.K.).

FIG. 1 *a*, Depth profile of formaldehyde photochemical production in the southwestern Sargasso Sea during March 1989 cruise; samples were irradiated in sunlight on the deck of the ship. Water depth was ~4600 m. *b*, As *a*, but for November 1989 cruise. *c*, DOM fluorescence (excitation wavelength 350 nm, emission wavelength 450 nm) during November 1989 cruise. QSU is quinine sulphate units; 1 QSU=1 part per 10⁹ quinine sulphate in 0.05M sulphuric acid.



photochemically reactive compounds by organisms and photo-bleaching of the fluorophores in the upper few metres of the water column are possible factors^{30,31}.

To estimate the ocean residence time of photochemically reactive DOC removed by the photochemical degradation pathway, several assumptions were made. We assumed that the measured photochemical production of LMW carbon species (12.9–26.7 nM-C h⁻¹, Table 1), based on short-term irradiations (4 h), is a good approximation for the range of photochemical degradation rates of photochemically reactive DOC during its passage through the sunlit layer of the ocean. Results of longer irradiations (up to 100 h) suggest that this is a reasonable assumption¹⁴. We further assumed that the ultraviolet degradable (or persulphate-oxidizable) portion of the DOC pool is fully degraded by sunlight into biologically oxidizable or volatile compounds^{12–16}. We found that <20% of the photochemically produced CO is oxidized microbially in those waters, and of this fraction >99% is oxidized to CO₂ (ref. 30). In our calculations, we normalized the measured photochemical production rates (Table 1) to the average concentration of this DOC fraction in the ocean, which we assumed to be 50 µM C. The latter assumption is based on a deep-sea concentration of 40–50 µM C and a surface water concentration of 60–80 µM C (refs 10, 17, 18). We also assumed an average ocean depth of 3,800 m and an average photochemically effective light penetration (1/*e*) depth of 5 m in the open ocean. The latter value was based on *in situ* irradiations and on model calculations of the penetration of the photochemically active wavelengths in open ocean water³². Action spectra showed that the wavelengths responsible for carbonyl photochemical production in the solar irradiance spectrum at the sea surface are predominantly in the ultraviolet-B region (280–320 nm)¹⁴. The resultant oceanic 1/*e* depth for photochemical production of carbonyl compounds and α-keto acids is ~3 m, which is similar to that obtained for nitrate photolysis in the sea³³. The 1/*e* depth for CO photochemical production, based on its action spectrum¹⁶, is ~6 m. In our calculations we used a weight-averaged 1/*e* depth of 5 m for carbonyl and CO photochemical production. Finally, we assumed 2–4 h of globally averaged, photochemically effective sunlight per day. This assumption was based on an average broad-band ultraviolet/visible light flux of 62 J cm⁻² d⁻¹ (or ~170 W h m⁻²) at the sea surface³³, which corresponds to ~2 h of solar noon irradiation in the subtropics (26° N) (K.M., X.Z. and G. Vaughan, unpublished results). The 2–4 h assumption is consistent with latitudinally dependent photochemical models^{34–37}. A more precise estimate of the globally averaged solar flux at the sea surface awaits resolution of discrepancies between model predicted and measured fluxes, especially in the ultraviolet-B region^{38–40}.

We calculated the ocean residence time of photochemically reactive DOC using a one-dimensional, steady-state model

described elsewhere⁴¹. In the calculations, we used an average deep water replacement time for the world ocean of 500 yr (ref. 42), and we assumed that the photochemical degradation rate (*R*) of photochemically reactive DOC is proportional to its concentration ([DOC])¹⁴, $R = -d[DOC]/dt = k[DOC]$, where *k* is the photochemical degradation rate constant for photochemically reactive DOC at the sea surface. The rate constant was estimated from measured photochemical production rates of LMW carbon species (12.9–26.7 nM-C per h, Table 1) and the above assumptions. The results indicated that ~12–48% of photochemically reactive DOC is degraded by sunlight per oceanic mixing cycle. This range corresponds to a range of *τ* of 1,000–4,200 yr.

Our estimated range of *τ* is less than the average apparent ¹⁴C age of 6,000 yr (ref. 2). The difference between these values is probably even larger because both estimates are conservative. The ultraviolet oxidation method used to convert DOC to CO₂ for ¹⁴C age determination also oxidizes some young DOC¹⁰. If a correction is made for this DOC, the apparent ¹⁴C age of the remaining, humic-rich fraction becomes older¹⁰. In our study we did not measure all possible photochemically formed LMW organic carbon species, such as formate and acetate, which have production rates similar to carbonyl compounds in other natural waters (K.M., X.Z. and G. Vaughan, unpublished results). Also, it is likely that CO₂ is produced during photochemical degradation of carboxyl-rich humic substances^{43,44}, and that photochemically degraded humic molecules are more susceptible to microbial uptake¹³. Thus, our measured photochemical production rates of LMW carbon species probably underestimate actual DOC photochemical degradation rates, possibly by a factor of two or more. A two-fold higher rate would yield a *τ* of 500–2,100 yr, which approaches a single oceanic mixing cycle⁴². This result suggests that the photochemical degradation pathway is even more important than calculated above and may be the rate-limiting step in the removal of biologically refractory DOC, such as humic substances, in the ocean.

The discrepancy between the average apparent ¹⁴C age of deep sea, photochemically reactive DOC and its photochemical residence time may be reconciled if there is an input of old, humic-rich DOC into deep water. One likely source for this DOC is sediment porewater, which has a 5–500 times higher DOC concentration and humic substance fluorescence than overlying water^{10,29,45–47}. A sediment impact can be clearly discerned in some deep-sea DOC profiles that extend to within 100 m of the sediment–water interface^{10,29,45,46}. Other direct and indirect evidence for a porewater input include fluorescence spectra^{48,49}, amino acid enantiomeric ratios^{48,49}, isotopic composition¹¹, NMR studies¹¹ and lignin residue analyses⁵⁰. At present, no flux measurements have been reported, but preliminary benthic chamber results from margin sediments suggest that this flux is important (K.M. and A. H. Devol, unpublished results).

Our results suggest that photochemical degradation has an important role in the cycling of biologically refractory DOC in the oceans. This role is likely to become more important in the future as depletion of atmospheric ozone increases the solar

ultraviolet-B flux³⁸. Also, our results support the idea that flux of DOC out of bottom sediments is an important component of the oceanic carbon cycle⁵¹ and so is worthy of further investigation. □

Received 7 June; accepted 12 July 1991.

1. Mopper, K. & Degens, E. T. in *The Global Carbon Cycle* (eds Bolin, B., Degens, E. T., Kempe, S. & Kettner, P.) 293–316 (Wiley, New York, 1979).
2. Williams, P. M. & Druffel, E. R. M. *Nature* **330**, 246–248 (1987).
3. Menzel, D. W. & Ryther, J. H. *Deep-Sea Res.* **15**, 327–337 (1968).
4. Barber, R. T. *Nature* **220**, 274–275 (1968).
5. Craig, H. J. *Geophys. Res.* **76**, 5078–5086 (1971).
6. Williams, P. M. & Carlucci, A. F. *Nature* **262**, 810–811 (1976).
7. Toggweiler, J. R. *Nature* **334**, 468 (1988).
8. Williams, P. M. & Druffel, E. R. M. *Oceanography* **1**, 14–17 (1988).
9. Toggweiler, J. R. in *Productivity of the Ocean: Present and Past* (eds Berger, W. H., Smetceck, V. S. & Wefer, G.) 65–83 (Wiley, New York, 1989).
10. Druffel, E. R. M., Williams, P. M. & Suzuki, Y. *Geophys. Res. Lett.* **16**, 991–994 (1989).
11. Malcolm, R. L. *Analyt. Chim. Acta* **232**, 19–30 (1990).
12. Kieber, D. J., McDaniel, J. & Mopper, K. *Nature* **341**, 637–639 (1989).
13. Geller, A. *Limnol. Oceanogr.* **31**, 755–764 (1986).
14. Kieber, R. J., Zhou, X. & Mopper, K. *Limnol. Oceanogr.* **35**, 1503–1515 (1990).
15. Redden, G. D. thesis, Oregon State Univ. (1983).
16. Valentine, R. L. & Zepp, R. G. *EOS* **71**, 1236 (1990); *Abstr. 201st National Meeting Am. chem. Soc. Div. Envir. Chem.* 524–527 (American Chemical Society, 1991).
17. Sugimura, Y. & Suzuki, Y. *Mar. Chem.* **24**, 105–131 (1988).
18. Suzuki, Y. & Tanoue, E. in *Ocean Margin Processes in Global Change* (eds Mantoura, R. F. C., Martin, J. M. & Wollast, R.) 18–23 (Dahlem Workshop, Wiley, Berlin, 1991).
19. Koike, I., Hara, S., Terauchi, K. & Kogure, K. *Nature* **345**, 242–244 (1990).
20. Toggweiler, J. R. *Nature* **345**, 203–204 (1990).
21. Mopper, K. & Zhou, X. *Science* **250**, 661–664 (1990).
22. Zafriou, O. C., Jousset-Dubien, J., Zepp, R. G. & Zika, R. G. *Envir. Sci. Technol.* **18**, 358A–371A (1984).
23. Kieber, R. J. & Mopper, K. *Envir. Sci. Technol.* **24**, 1477–1481 (1990).
24. Mopper, K. & Stahovec, W. L. *Mar. Chem.* **19**, 305–321 (1986).
25. Kieber, D. J. & Mopper, K. *Mar. Chem.* **21**, 135–149 (1987).
26. Butler, J. H., Pequegnat, J. E., Gordon, L. I. & Jones, R. D. *Estuar. Coast. Shelf Sci.* **27**, 181–203 (1983).
27. Jones, R. D. *Deep-Sea Res.* **38**, 625–635 (1991).
28. De Haan, H. & De Boer, T. *Water Res.* **21**, 731–734 (1987); Stewart, A. J. & Wetzel, R. G. *Limnol. Oceanogr.* **25**, 559–564 (1980).

29. Chen, R. F. & Bada, J. L. *Geophys. Res. Lett.* **16**, 687–690 (1989).
30. Blough, N. V. & Zepp, R. G. (eds) *Effects of Solar Ultraviolet Radiation of Biogeochemical Dynamics in Aquatic Environments* (Woods Hole Oceanographic Institution Techn. Rep. 90-09, Massachusetts, 1990).
31. Kouassi, A. M., Zika, R. G. & Plane, J. M. C. *Neth. J. Sea Res.* **27**, 33–41 (1990).
32. Baker, K. S., Smith, R. C. & Green, A. E. S. in *The Role of Solar Ultraviolet Radiation in Marine Ecosystems* (ed. Calkins, J.) 79–91 (Plenum, New York, 1982).
33. Zafriou, O. C. & True, M. B. *Mar. Chem.* **8**, 33–42 (1979).
34. Dahlback, A., Henriksen, T., Larsen, S. H. & Starnes, K. *Photochem. Photobiol.* **49**, 621 (1989).
35. Cutchis, P. *Science* **184**, 13–19 (1974).
36. Zepp, R. G., Hoigné, J. & Bader, H. *Envir. Sci. Technol.* **21**, 443–450 (1987).
37. Zepp, R. G. & Cline, D. M. *Envir. Sci. Technol.* **11**, 359–366 (1977).
38. Spinhirne, J. D. & Green, A. E. S. *Atmos. Envir.* **12**, 2449–2454 (1978).
39. Smith, R. C. & Baker, K. S. *Oceanography* **2**, 4–10 (1989).
40. Starnes, K., Slusser, J., Bowen, M., Booth, C. & Lucas, T. *Geophys. Res. Lett.* **17**, 2181–2184 (1990).
41. Broecker, W. S. & Peng, T.-H. *Tracers in the Sea* (Eldigio, New York, 1982).
42. Stuiver, M., Quay, P. D. & Ostlund, H. G. *Science* **219**, 849–854 (1983).
43. Mathews, R. W. *Radiat. Res.* **83**, 27–41 (1980).
44. Wang, W. H., Beyerle-Pfäur, R. & Lay, J. P. *Chemosphere* **17**, 1197–1204 (1988).
45. Williams, P. M., Carlucci, A. F. & Olson, R. *Oceanol. Acta* **3**, 471–476 (1980).
46. Starikova, N. D. *Oceanol.* **10**, 796–807 (1970).
47. Devol, A. H. & Ahmed, S. I. *Nature* **291**, 407–408 (1981).
48. Mopper, K. in *Proc. MarChem '91 Workshop* (Office of Naval Research, Steamboat Springs, Colorado, 1991).
49. Mopper, K. & Park, Y. C. in *Background Papers Workshop on Measurement of Dissolved Organic Carbon and Nitrogen in Natural Waters* (NSF/NOAA/DOE, Seattle, 1991).
50. Meyers-Schulte, K. J. & Hedges, J. I. *Nature* **321**, 61–63 (1986).
51. Hedges, J. I. *Mar. Chem.* (in the press).

ACKNOWLEDGEMENTS. We thank J. Bauer, E. R. M. Druffel, C. Lee, J. I. Hedges, J. R. Toggweiler and P. M. Williams for comments on the manuscript. J. A. McDaniel, G. Vaughan and J. A. Amador for technical assistance during irradiation experiments, and the captain and crew of the RV *Columbus Iselin* for assistance during sampling. Financial support was provided by the NSF's Chemical Oceanography Program and the Office of Naval Research's Ocean Chemistry Program.

Increased mantle melting beneath Snaefellsjökull volcano during Late Pleistocene deglaciation

Björn S. Hardarson & J. Godfrey Fitton

Department of Geology and Geophysics, University of Edinburgh,
West Mains Road, Edinburgh EH9 3JW, UK

BASALTIC magmatism results when upwelling mantle crosses the peridotite solidus. The greater the overstep of the solidus, the greater the degree of mantle melting and the larger the volume of magma produced. In Iceland most of the active volcanism occurs along the central spreading axis, although some occurs in isolated off-axis volcanoes such as Snaefellsjökull. Because the mantle beneath Snaefellsjökull is not upwelling as vigorously as that beneath the spreading axis, it oversteps its solidus by a much smaller amount and consequently the degree of melting is less. The composition of the resulting magma will be much more sensitive to small perturbations in the amount of upwelling than will the mantle beneath the ridge axis. We show here that the reduction of pressure in the mantle caused by the unloading of ice was sufficient to affect magma composition. Unloading was accompanied by a clear shift towards less undersaturated magma compositions, which reflect a transient increase of about 0.5% in the degree of mantle melting, coupled with a decrease in depth of melting.

The volume and composition of basaltic magma can be related quantitatively to the amount of lithospheric extension, the thickness of the lithosphere before extension, and the potential temperature of the asthenospheric mantle¹. In Iceland, anomalously hot mantle upwelling beneath the mid-Atlantic ridge results in greater volumes of magma and hence greater crustal thickness than elsewhere along the ridge. The largest degrees of melting occur beneath the axis of the spreading centre and the resulting

magma is intruded or erupted in a narrow, volcanically active belt running across Iceland. These large-volume magmas are tholeiitic in character. Off-axis volcanoes such as Snaefellsjökull tend to erupt transitional to mildly alkaline magmas.

The downward deflection of the lithosphere due to loading of an ice cap causes outflow of the asthenospheric material beneath. When the ice retreats the crust is left depressed and gradually rebounds as the viscous asthenosphere flows back. If the asthenosphere is partially molten, such isostatic movements could affect the amount of melt present by suppressing melting during loading and subsidence, and enhancing it during unloading and rebound. The effect on magma composition is likely to be very small and might only manifest itself in situations where the degree of melting is very low. The Snaefellsjökull volcanic system, located 150 km from the spreading axis, on the Snaefellsnes Peninsula in western Iceland (64.8° N, 23.8° W), is one place where such effects might be observed.

The beginning of the Quaternary era in Iceland is fixed at the base of the Mammoth magnetic event at 3.1 Myr ago, at which time the first indications of the onset of the ice age in Iceland are found². The remains of numerous glacial stages are known with interglacial stages in between. The oldest exposed volcanic rocks from the Snaefellsjökull system are thought to be ~700,000 years old³ and from that time to the present at least four glacial stages are known², including the last one which began ~75,000–70,000 years ago and was at its maximum ~18,000 years ago. Melting of the ice cap began ~15,000 years ago and was interrupted by two brief returns to near-glacial conditions at 12,500–12,000 years and 11,000–10,000 years (corresponding to the older and younger Dryas, respectively). Uplift of Iceland took place before the end of the Pleistocene, and marine sediments are now, for example, found up to ~170 m above sea level in Bulandshöfði in the northern Snaefellsnes peninsula.

Basaltic rock samples collected from Snaefellsjökull can be divided into three chronological groups on the basis of field relations. The oldest rocks were erupted during subglacial and interglacial stages at and following the Matuyama–Brunhes

transition. These rocks form the early glacial (EG) group. The late glacial (LG) group comprises lavas and hyaloclastite deposits erupted shortly before, between, or during the two short-lived Dryas glacial advances. The post-glacial (PG) group of lavas was erupted after the final retreat of the ice. The EG rocks were mostly erupted at a time when glacial conditions were fully established over the area and the LG rocks soon after the melting of the Iceland ice cap.

A comparison of the composition of basaltic rocks ($\text{MgO} > 4$ wt%) from each of these three groups is given in Table 1. It is clear from these data that EG and PG rocks are very similar to each other but that the LG rocks, erupted at a time of rapid climatic change⁵, are significantly different. The abundances of the more incompatible elements (K, Rb, Ba, Nb, Th, Pb, La, Ce, Nd) are lower by up to a factor of two in these rocks than in the other two groups. These differences are clearly displayed in plots involving ratios of elements of contrasted compatibility (for example Fig. 1). The ratios Ce/Y and Zr/Nb are insensitive to moderate degrees of low-pressure crystal fractionation involving mineral phases likely to crystallize from basaltic magma. The variation shown in Fig. 1 must, therefore, reflect differences in the degree of mantle melting and/or differences in source composition.

The effects of variation of source composition and degree of melting are shown in Fig. 1 using fractional melting curves calculated for depleted and primitive mantle compositions and for spinel- and garnet-lherzolite mineralogies. The Snaefellsjökull data fall between the garnet- and spinel-lherzolite curves, consistent with derivation of the magmas from a melt column extending across the garnet-spinel transition. A similar conclusion was reached by McKenzie and O'Nions⁷ on the basis of rare-earth element data from Icelandic basalts.

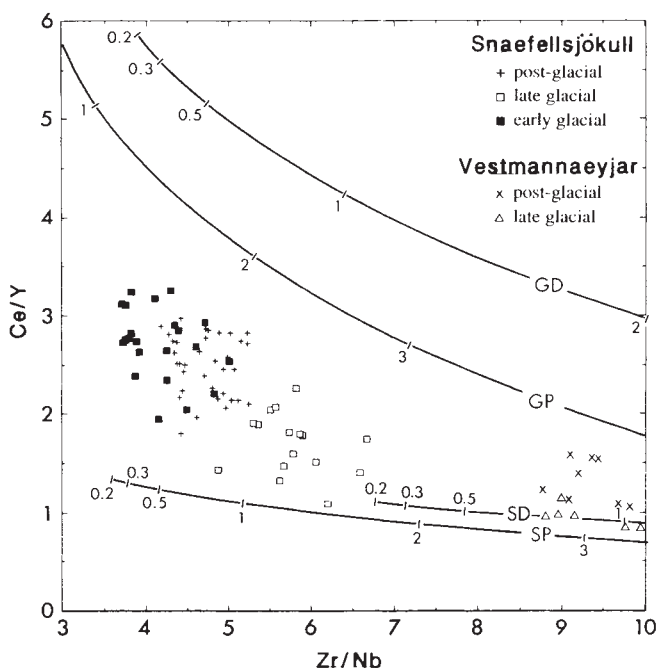


FIG. 1 The Ce/Y ratio against Zr/Nb for volcanic rocks from Snaefellsjökull and Vestmannaeyjar. The lines are non-modal fractional melting curves for plausible mantle compositions: GD, depleted garnet lherzolite; GP, primitive garnet lherzolite; SD, depleted spinel lherzolite; SP, primitive spinel lherzolite. Numbers against the curves are percentages of melt. Elemental ratios in mid-ocean-ridge basalt and chondritic meteorites⁶ were taken to represent depleted and primitive mantle sources, respectively. Mantle mineralogies and partition coefficient (D) data were taken from ref. 7. D_Y was taken to be equal to D_{Hf} ; D_{Nb} was assumed to be equal to D_{La} because La/Nb is virtually constant in ocean-island basalts. Analytical reproducibility (2σ): Ce/Y, $\pm 10\%$; Zr/Nb, $\pm 1\%$.

Whatever combination of curves is used to model the Snaefellsjökull data, it is clear from Fig. 1 that the magmas were produced by less than 2% melting. The shift in magma composition shown by the LG group can best be explained by a transient increase in degree of melting by a little under 1% for a primitive mantle source, coupled with a decrease in depth of melting (into the spinel-lherzolite field). For a depleted mantle source the required increase in degree of melting is $\sim 0.5\%$.

Mantle melting beneath Iceland will largely occur within the spinel-lherzolite field but with a contribution from small-degree melts generated from garnet-lherzolite at greater depths⁷. These small-degree melts account for only a small proportion of the total volume of magma but will contribute a large proportion of the more incompatible elements. This accounts for the apparent influence of garnet-lherzolite in the generation of the Snaefellsjökull magmas (Fig. 1). These small melt fractions will rise to the surface with a velocity of only a few millimetres per year⁸ and will not therefore be able to contribute to the transient increase in degree of melting. All the excess melt that reaches the surface over this short period will have been generated at shallow depths within the spinel-lherzolite field. This accounts for the shift towards the spinel-lherzolite melting curves shown by the LG data in Fig. 1.

If isostatic movements affect melting processes by suppressing the degree of melting during subsidence and enhancing it during crustal rebound, then such a process could explain the chemical variation seen in Fig. 1. The EG and PG volcanic rocks were erupted at times when the crust was fully depressed and almost fully recovered, respectively. These are the most undersaturated rocks on Snaefellsjökull and represent small degrees of mantle melting. The intervening LG rocks were erupted just after unloading, during the most rapid phase of crustal rebound, and

TABLE 1 Mean composition of basaltic rocks ($\text{MgO} > 4$ wt%) from Snaefellsjökull

	Early Glacial		Late Glacial		Post-glacial	
	Mean	1σ	Mean	1σ	Mean	1σ
	$n=22$		$n=17$		$n=35$	
	(wt%)					
SiO ₂	47.14	1.03	46.62	0.57	46.44	1.31
Al ₂ O ₃	14.85	1.23	15.79	0.68	14.97	0.63
Fe ₂ O ₃	12.48	0.96	12.14	0.63	14.16	1.20
MgO	7.73	2.87	7.54	1.16	6.03	1.96
CaO	10.30	1.30	12.13	0.50	9.99	1.06
Na ₂ O	2.72	0.61	2.61	0.34	3.13	0.58
K ₂ O	1.09	0.30	0.60	0.12	1.11	0.30
TiO ₂	2.83	0.55	2.39	0.25	3.47	0.52
MnO	0.21	0.02	0.19	0.02	0.24	0.03
P ₂ O ₅	0.62	0.22	0.40	0.25	0.83	0.31
Total	99.97		100.41		100.37	
	Parts per million					
Ni	107	88	83	38	49	61
Cr	349	332	171	144	104	207
V	318	44	339	24	378	77
Sc	31.5	5.4	36.6	3.5	31.6	4.9
Cu	58	26	81	18	46	23
Zn	98	18	80	8	111	16
Sr	491	60	430	47	497	35
Rb	22.3	7.3	11.1	2.8	20.8	6.9
Zr	193	49	137	23	236	63
Nb	47.4	12.8	24.0	4.5	50.4	14.6
Ba	413	100	241	68	507	136
Pb	1.7	1.2	1.0	0.8	1.7	0.7
Th	2.8	1.0	1.6	0.9	2.6	1.0
La	33.6	10.3	15.8	5.8	39.1	13.5
Ce	74.7	19.8	39.3	8.5	85.9	27.4
Nd	37.2	8.6	23.0	3.9	44.2	14.2
Y	27.6	6.2	22.9	3.0	34.0	8.4

Based on unpublished X-ray fluorescence data. Details of the method and analytical precision are given in ref. 4.

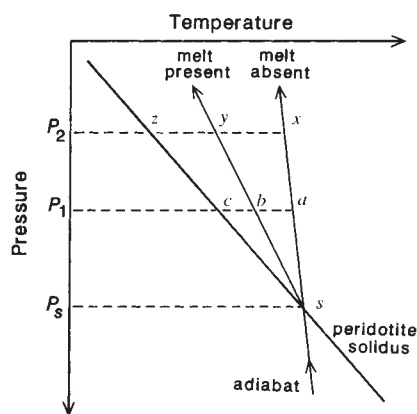


FIG. 2 Pressure-temperature diagram illustrating schematically the effect of changes in pressure on the temperature overstep on the solidus. Mantle rising along an adiabat intersects the solidus at point *s* at pressure P_s , and melt is generated between *s* and *b*. Reducing the pressure from P_1 to P_2 extends the melt-present curve to *y* with an increased temperature overstep and a larger degree of melting.

are the products of larger degrees of melting. Recent evidence suggests that the isostatic rebound may have happened very rapidly^{9,10}, possibly in less than 700 years⁹.

To demonstrate that the observed shift in magma composition could be related to the unloading of ice, we must show that the resulting decompression of the mantle was sufficient to account for the required increase in the degree of mantle melting. The effect of a small decrease in pressure on the mantle is illustrated in Fig. 2. Mantle upwells adiabatically until it crosses its solidus at pressure P_s . If the mantle could upwell without melting it would continue along an adiabat to point *a* at pressure P_1 with a temperature overstep above the solidus of $\Delta\Theta_1 (=T_a - T_c)$. The effect of melting would be to reduce the temperature, and the mantle plus melt would rise along an isentropic curve to *b* with a temperature overstep of $\Delta T_1 (=T_b - T_c)$. A further reduction in pressure to P_2 would result in a temperature overstep of $\Delta\Theta_2 (=T_x - T_z)$ in the absence of melting and $\Delta T_2 (=T_y - T_z)$ after melting. The respective increases in temperature overstep would be $\Delta\Theta (= \Delta\Theta_2 - \Delta\Theta_1)$ and $\Delta T (= \Delta T_2 - \Delta T_1)$.

The melt-absent temperature overstep $\Delta\Theta$ can be calculated from the pressure drop $\Delta P (=P_1 - P_2)$ and the gradient of the solidus (dT/dP). The maximum thickness of the Icelandic icecap was ~2 km (ref. 11) which would, on melting, reduce the pressure on the underlying mantle by 0.2 kbar. Experimental determinations of the peridotite solidus (compiled by McKenzie and Bickle¹) give $dT/dP = 13.6^\circ\text{C kbar}^{-1}$. The overstep $\Delta\Theta$ resulting from ice unloading would therefore be 2.4°C , allowing for the effects of adiabatic decompression ($1.4^\circ\text{C kbar}^{-1}$). The calculation of the increase in temperature overstep after melting (ΔT) is less straightforward⁸ and was calculated using Watson and McKenzie's parameterization (ref. 12, equation 15) which relates ΔT to $\Delta\Theta$ and P for an entropy change $\Delta S = 250 \text{ J kg}^{-1}^\circ\text{C}^{-1}$. This gave a value of 1.14°C for ΔT at a pressure of 10 kbar.

The increase in temperature overstep after melting (ΔT) caused by unloading can be related directly to an increase in melt fraction (ΔX) using McKenzie and Bickle's¹ parameterization of experimental data. Using their equations 18–21 gives an increase of 0.4% for melt fractions close to zero at a pressure of 10 kbar. McKenzie and Bickle¹ fitted a smooth curve to the experimentally determined melt fractions, although there is some evidence for near-invariant (eutectic) behaviour close to the solidus¹³. The increase in melt fraction produced by eutectic melting will be larger than that calculated from McKenzie and Bickle's parameterization, because all of the excess heat due to the temperature overstep is converted into melt. In this case the melt-present line on Fig. 1 will run along the solidus ($\Delta T = 0$). The increase in melt fraction resulting from eutectic melting can easily be calculated from

$$\Delta X = C_p \Delta\Theta / T \Delta S$$

where C_p is the specific heat of the mantle at constant pressure, ΔS is the entropy change on melting and T is the absolute temperature. The effects of the pressure change and the

difference in the specific heats of mantle and melt are small and are neglected. Taking $C_p = 10^3 \text{ J kg}^{-1}^\circ\text{C}^{-1}$, $\Delta S = 250 \text{ J kg}^{-1}^\circ\text{C}^{-1}$, and a temperature of 1,600 K gives a melt increase of 0.6%. This is a maximum value for the increase in melting. Thus, unloading of the ice cap could have increased the melt fraction in the underlying mantle by between 0.4 and 0.6%, a figure comparable with that required by the chemical data from Snaefellsjökull assuming a depleted mantle source. A source depleted in incompatible elements with respect to primitive mantle is implied by isotope data from Snaefellsjökull ($^{87}\text{Sr}/^{86}\text{Sr} = 0.70328\text{--}0.70337$; $^{143}\text{Nd}/^{144}\text{Nd} = 0.51296\text{--}0.51304$; B.S.H., unpublished data).

Differences in degree of melting of ~0.5% are not likely to affect the composition of tholeiitic magmas generated in the rift zone where degrees of melting are high. The eruptive products of Krafla, for example, show no compositional variations over the late stages of the ice age¹⁴, although Sigvaldason and Steinthorsson¹⁵ have argued that pressure release during isostatic rebound was responsible for the eruption of large volumes of depleted olivine tholeiite in central Iceland. Where the degree of melting is small, however, as in the Snaefellsjökull volcanic system, the small variation in degree of melting accompanying the unloading of ice is large enough to affect magma composition.

Some support for the conclusions reached here is provided by late glacial¹⁶ and post-glacial basaltic rocks from Vestmannaeyjar (Westman Islands), located at the tip of the southward-propagating eastern rift zone, which make up the only other alkaline volcanic system in Iceland. Unfortunately, early glacial rocks are not exposed on these islands. Our data for rocks from Vestmannaeyjar (Fig. 1) show that the magmas were produced by larger degrees of melting than were those from Snaefellsjökull and as a consequence there is no evidence for variation in degree of melting (the two groups have comparable Zr/Nb). There is, however, a clear shift towards a more garnet-rich source (high Ce/Y) for the post-glacial lavas, comparable to that seen in lavas of the same age from Snaefellsjökull. □

Received 13 May; accepted 18 July 1991.

- McKenzie, D. & Bickle, M. J. *J. Petrol.* **29**, 625–679 (1988).
- Saemundsson, K. *Jökull* **29**, 7–28 (1980).
- Johannesson, H., Flores, R. M. & Jonsson, J. *Jökull* **31**, 23–30 (1981).
- Fitton, J. G. & Dunlop, H. M. *Earth planet. Sci. Lett.* **72**, 23–38 (1985).
- Fairbanks, R. G. *Nature* **342**, 637–642 (1989).
- Hofmann, A. W. *Earth planet. Sci. Lett.* **90**, 297–314 (1988).
- McKenzie, D. & O'Nions, R. K. *J. Petrol.* (in the press).
- McKenzie, D. P. *J. Petrol.* **25**, 713–765 (1984).
- Thors, K. & Boulton, G. S. *Marine Geol.* **98**, 99–112 (1991).
- Thors, K. & Helgadóttir, G. in *Environmental Change in Iceland* (eds Maizels, J. & Caseldine, C.) (in the press).
- Einarsson, D. *Geology, Story of Rock and Country* (in Icelandic) 355 (Mal & Menning, Reykjavik, 1968).
- Watson, S. & McKenzie, D. *J. Petrol.* **32**, 501–537 (1991).
- Mysen, B. O. & Kushiro, I. *Am. Miner.* **62**, 843–865 (1977).
- Nicholson, H. thesis, Univ. of Edinburgh (1990).
- Sigvaldason, G. E. & Steinthorsson, S. *Geodynamics of Iceland and the North Atlantic Area* (ed. Kristjánsson, L.) 155–164 (Reidel, Dordrecht, 1974).
- Jakobsson, S. P. *Acta Nat. Islandica* **26**, 103 (1979).

ACKNOWLEDGEMENTS. We thank M. Bickle, D. Latin, D. McKenzie and S. Steinthorsson for their comments.

Binding of synaptotagmin to the α -latrotoxin receptor implicates both in synaptic vesicle exocytosis

Alexander G. Petrenko*†, Mark S. Perin*, Bazbek A. Davletov†, Yuri A. Ushkaryov*, Martin Geppert* & Thomas C. Südhof*‡

* Howard Hughes Medical Institute and Department of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, Texas 75235, USA

† Shemyakin Institute of Bioorganic Chemistry, USSR Academy of Sciences, Moscow 117871, USSR

‡ To whom correspondence should be addressed

A VERTEBRATE neurotoxin, α -latrotoxin, from black widow spider venom causes synaptic vesicle exocytosis and neurotransmitter release from presynaptic nerve terminals¹⁻⁴. Although the mechanism of action of α -latrotoxin is not known, it does require binding of α -latrotoxin to a high-affinity receptor on the presynaptic plasma membrane⁵. The α -latrotoxin receptor seems to be exclusively at the presynaptic plasmamembrane⁶. Here we report that the α -latrotoxin receptor specifically binds to a synaptic vesicle protein, synaptotagmin, and modulates its phosphorylation. Synaptotagmin is a synaptic vesicle-specific membrane protein that binds negatively charged phospholipids and contains two copies of a putative Ca^{2+} -binding domain from protein kinase C (the C2-domain), suggesting a regulatory role in synaptic vesicle fusion^{7,8}. Our findings suggest that a physiological role of the α -latrotoxin receptor may be the docking of synaptic vesicles at the active zone. The direct interaction of the α -latrotoxin receptor with a synaptic vesicle protein also suggests a mechanism of action for this toxin in causing neurotransmitter release.

To investigate potential interactions of the α -latrotoxin receptor with other synaptic proteins, the receptor was purified from solubilized bovine brain membrane proteins by affinity chromatography on immobilized α -latrotoxin^{9,10}. Using a two-stage KCl-gradient, three major polypeptides (relative molecular masses 79,000, 65,000 and 43,000) with no α -latrotoxin-binding activity eluted from the column during the first stage of the gradient, whereas the actual α -latrotoxin receptor, as defined by α -latrotoxin binding, eluted during the second stage¹⁰ (Fig. 1). The 65K protein stained poorly with Coomassie blue but

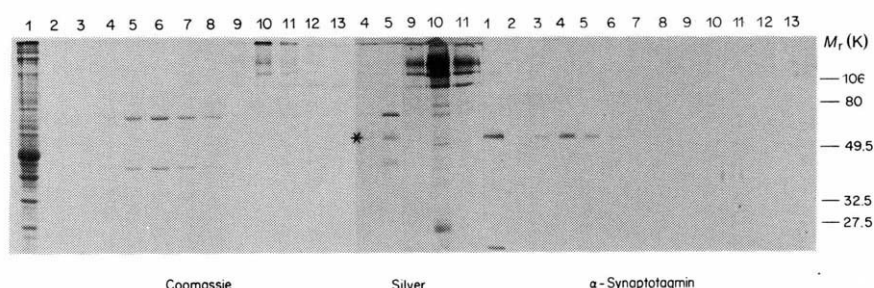
silver staining and amino-acid sequencing suggested that these three polypeptides and the major protein components of the α -latrotoxin receptor were eluted in comparable quantities (Fig. 1). Partial sequences of the 79K and 43K proteins revealed no significant homology to proteins in the databanks but the sequences of the 65K protein were found to be identical to those of synaptotagmin (data not shown)^{7,8}. Identification of the 65K protein as synaptotagmin was confirmed by immunoblotting, which demonstrated the specific elution of synaptotagmin from the receptor/ α -latrotoxin column (Fig. 1).

In protein kinase C, the C2-domain is thought to represent a Ca^{2+} -binding domain¹¹, suggesting that synaptotagmin is also a Ca^{2+} -binding protein. Recombinant synaptotagmin also avidly binds negatively charged phospholipids⁷. Because of these properties synaptotagmin was postulated to have a role in the docking and fusion of synaptic vesicles. A direct interaction of synaptotagmin with the receptor for α -latrotoxin, a neurotoxin that causes massive neurotransmitter release, would suggest that this receptor has a physiological role in the docking and fusion of synaptic vesicles. Thus we investigated the specificity of this interaction.

Membrane proteins from partially purified rat synaptic vesicles were chromatographed on columns containing either the bovine α -latrotoxin receptor bound to immobilized α -latrotoxin or immobilized α -latrotoxin alone without the receptor. In addition, several identically prepared control columns containing immobilized calmodulin, rab3A and bovine serum albumin were used. Synaptotagmin was only bound to and eluted from the column containing prebound receptor, but not from the α -latrotoxin column without receptor (Fig. 2a) or from the other control columns (data not shown). Synaptotagmin was the major protein eluted from the α -latrotoxin receptor column with only trace amounts of other proteins present. Therefore synaptotagmin can be affinity-purified on the α -latrotoxin receptor in a one-step procedure in a manner that is specific for the α -latrotoxin receptor but independent of the 79K and 43K proteins, suggesting a direct interaction of synaptotagmin with the α -latrotoxin receptor.

To determine whether the affinity-purified synaptotagmin was indeed from the rat synaptic vesicles and was not carried over from the previous receptor purification from bovine brain, we used an antibody that recognizes rat but not bovine synaptotagmin¹². This antibody reacted only with the synaptotagmin eluted after rat synaptic vesicles had been loaded onto the column containing bovine receptor but not with the bovine synaptotagmin loaded during the first elution (Fig. 2b). Furthermore, immunoblotting with a panel of antibodies against >10 different synaptic vesicle proteins¹³ demonstrated that only synaptotag-

FIG. 1 Copurification of synaptotagmin with the α -latrotoxin receptor on immobilized α -latrotoxin. Solubilized membrane proteins from bovine brain (lane 1) were loaded onto an immobilized α -latrotoxin column and eluted with a two-stage salt gradient¹⁰. Three major proteins of 79K, 65K (synaptotagmin) and 43K elute during the first stage of the gradient whereas the α -latrotoxin receptor with major polypeptide components of 200K, 160K and 29K elutes during the second stage of the gradient. Synaptotagmin is almost invisible by Coomassie-blue staining (left) but is clearly visible by silver staining (asterisk, middle panel) and by immunoblotting (right) and was observed in previous preparations as a 35K proteolytic breakdown product¹⁰. Similarly, the 29K component of the α -latrotoxin receptor which is not a proteolytic breakdown product of the 200K and 160K polypeptides as indicated by amino-acid sequencing (A.G.P. *et al.*, unpublished results) also stains poorly with Coomassie blue but well with silver. Synaptotagmin was identified by amino-acid sequencing in a molar yield comparable to that of the 79K and 43K polypeptides (data not shown). The 120K protein found throughout the gradient may represent α -latrotoxin bleeding from the column.



METHODS. Bovine brain membranes were solubilized in 2% (v/v) Triton X-100 (lane 1), and loaded onto an immobilized α -latrotoxin column¹⁰. The column was eluted first with a 0.13–0.6 M KCl gradient containing 2 mM Ca^{2+} (lanes 2–8) and then with a 0.6 M KCl, 2 mM Ca^{2+} to 1.0 M KCl, 10 mM EDTA gradient (lanes 9–13). Samples were analysed on a 9% SDS-polyacrylamide gel followed by Coomassie blue staining (left panel), silver staining (middle panel) and immunoblotting with a polyclonal antibody against synaptotagmin (right panel)¹².

min was bound to the α -latrotoxin receptor column (Fig. 2b and data not shown).

Thus synaptotagmin binding to the immobilized α -latrotoxin column is specific for the α -latrotoxin receptor on the column matrix and for synaptotagmin as the synaptic vesicle protein. The converse affinity chromatography experiment was also done by using a column containing recombinant synaptotagmin produced in bacteria. Chromatography of solubilized bovine brain membrane proteins on the immobilized synaptotagmin resulted in a strong enrichment particularly of the α -latrotoxin receptor subunit of relative molecular mass (M_r) 200,000 (200K), again showing a specific interaction (data not shown).

To assay independently for the interaction of the α -latrotoxin receptor with synaptotagmin, we investigated the phosphorylation of synaptotagmin. Purified synaptotagmin was efficiently phosphorylated *in vitro* on the addition of [32 P]ATP or [32 P]GTP and this phosphorylation was potently inhibited by low concentrations of the α -latrotoxin receptor (Fig. 3). As synaptotagmin has no homology to the catalytic domains of protein kinases⁷, it is presumably phosphorylated by a trace contamination of an endogenous protein kinase. This protein kinase uses 1 μ M GTP as efficiently as 1 μ M ATP, phosphorylates synaptotagmin exclusively on threonine, and is stimulated by Ca^{2+} although it is not absolutely Ca^{2+} -dependent (Fig. 3, and data not shown).

Purified α -latrotoxin receptor inhibited synaptotagmin phosphorylation at low concentrations in a dose-dependent manner (Fig. 3a). Inhibition of synaptotagmin phosphorylation by purified α -latrotoxin receptor was abolished by boiling. Inhibition occurred only when the receptor was added before the phosphorylation reaction but not when added afterwards, indicating that the receptor does not contain phosphatase activity (Fig. 3). Synaptotagmin contains a hypersensitive proteolytic site that is located immediately C terminal to its transmembrane region and cleaves synaptotagmin into cytoplasmic

and membrane-bound fragments (see bottom of Fig. 3a)¹². Proteolysis of phosphorylated synaptotagmin demonstrates that synaptotagmin is phosphorylated exclusively on its cytoplasmic sequences and does not require membrane anchoring (Fig. 3a). The inhibition of synaptotagmin phosphorylation by the α -latrotoxin receptor suggests that synaptotagmin and the receptor interact through the cytoplasmic domains of synaptotagmin.

To investigate further the phosphorylation-inhibitory activity of the purified α -latrotoxin receptor, the receptor was fractionated on sucrose gradients either free or bound to α -latrotoxin. Immunoblotting of the fractions with antibodies against the 200K and 160K receptor proteins revealed a large shift in their apparent M_r as a function of α -latrotoxin binding (Fig. 3b). Assays of the synaptotagmin-phosphorylation inhibitory activity of the fractions demonstrated that the activity comigrated with the free 200K and 160K polypeptides in the absence of α -latrotoxin and was shifted to high M_r together with the receptor in the presence of α -latrotoxin. Thus the receptor protein is directly inhibitory to synaptotagmin phosphorylation and interacts with synaptotagmin independent of its occupancy by α -latrotoxin, although α -latrotoxin may modulate this interaction.

No *in vivo* phosphorylation of synaptotagmin has been observed¹⁴ and it is unclear if synaptotagmin phosphorylation has a physiological regulatory role. Little phosphorylation of synaptotagmin is apparent in intact synaptic vesicles but synaptotagmin is efficiently phosphorylated after solubilization of synaptic vesicles (Fig. 4), suggesting activation of the kinase phosphorylating synaptotagmin by detergents. Phosphorylation of synaptotagmin in synaptic vesicles is also completely inhibited by low concentrations of α -latrotoxin receptor (Fig. 4). The inhibitory action of the α -latrotoxin receptor is highly selective. It does not inhibit the phosphorylation of other proteins in the partially purified synaptic vesicles, suggesting that the α -latrotoxin receptor is not a general protein kinase inhibitor but that it selectively modulates synaptotagmin phosphorylation.

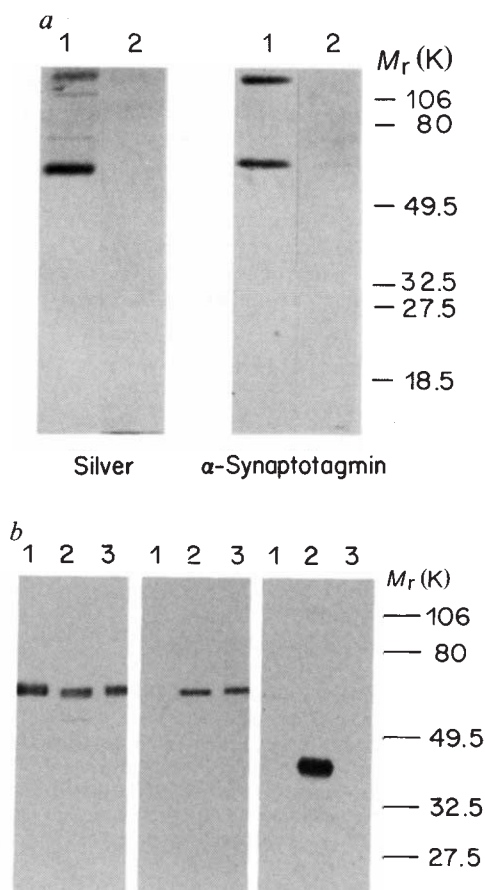
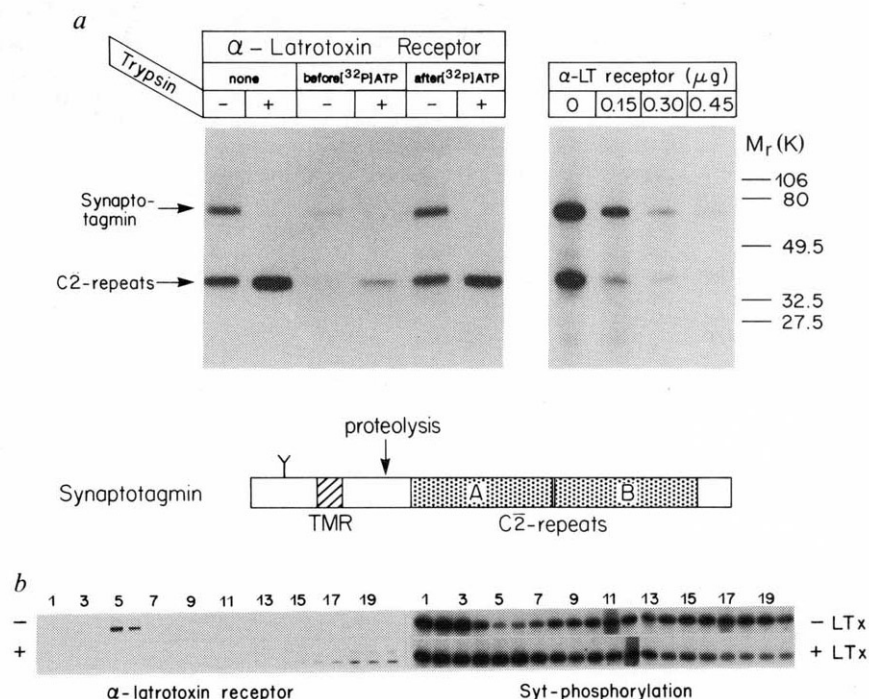


FIG. 2 Specificity of the binding of synaptotagmin to the α -latrotoxin receptor. *a*, Solubilized proteins from partially purified rat brain synaptic vesicles were chromatographed on columns containing either bovine α -latrotoxin receptor attached to α -latrotoxin but devoid of the 79K and 43K polypeptides (lane 1) or immobilized α -latrotoxin only (lane 2). Eluted proteins were analysed by SDS-PAGE and silver staining (left panel) or immunoblotting (right panel), demonstrating that synaptotagmin is the major protein purified. The 130K protein eluted is a synaptotagmin dimer¹². *b*, Immunoblotting analysis of the specificity of binding of rat synaptotagmin to a bovine α -latrotoxin receptor column. Lane 1 contains the polypeptides eluting during the first stage of the gradient of the α -latrotoxin receptor chromatography from bovine brain (fraction 5 in Fig. 1), lane 2 partially purified synaptic vesicles and lane 3 the proteins eluted from the α -latrotoxin receptor column that had been loaded with synaptic vesicle proteins. Fractions were analysed with antibodies that react with bovine and rat synaptotagmin (left panel), or only with rat but not bovine synaptotagmin (middle panel), or an antibody against another synaptic vesicle protein, synaptophysin (right panel). In parallel blots, samples were also analysed for synapsins Ia, Ib, IIa and IIb, synaptoporin, synaptobrevins I and II, rab3A and 3B, SV2 and p29 with results similar to the right panel.

METHODS. Partially purified rat brain synaptic vesicles¹² were solubilized in 2% (v/v) Triton X-100 and loaded onto either an immobilized α -latrotoxin column or an immobilized α -latrotoxin column containing the bovine α -latrotoxin receptor but not the accessory proteins (after the first stage of the gradient of Fig. 1). Further control columns used contained covalently coupled bovine serum albumin, calmodulin or rab3A (data not shown). Columns were equilibrated in 0.13M KCl, 2 mM Ca^{2+} and buffer and eluted as described in the legend to Fig. 1. Fractions were analysed by SDS-PAGE and silver staining or immunoblotting¹². References to the different synaptic vesicle proteins and antibodies against them can be found in refs 8, 13, 19 and 20.

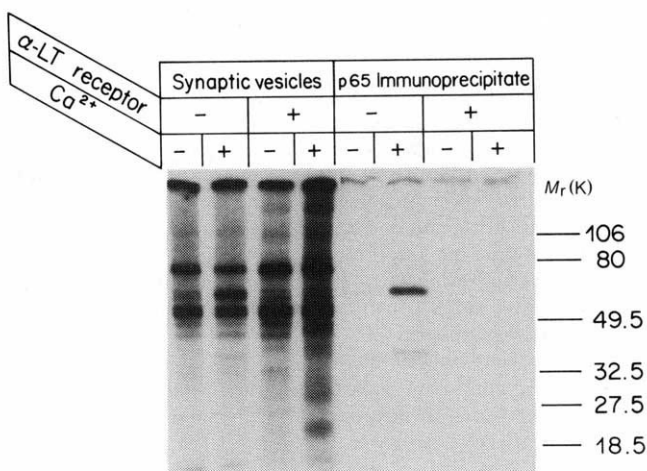
FIG. 3 Inhibition of synaptotagmin phosphorylation by the α -latrotoxin receptor. *a*, Synaptotagmin was phosphorylated by an endogenous protein kinase with γ [32 P]-ATP, with purified α -latrotoxin receptor added either before or after the phosphorylation reaction to demonstrate dependence of the phosphorylation inhibition on earlier addition of the receptor. Right, the inhibition of synaptotagmin phosphorylation by the α -latrotoxin receptor was titrated with increasing amounts of receptor, suggesting stoichiometric binding. Bottom, the structure of synaptotagmin is shown with the relative positions of the C2-repeats, the hypersensitive proteolytic site, and the transmembrane region (TMR)¹². The synaptotagmin preparation used here is partially cleaved at the hypersensitive proteolytic site by endogenous proteolysis and was completely cleaved by trypsin, demonstrating that the phosphorylation occurs on the cytoplasmic domain of synaptotagmin in the vicinity of the C2-repeats. *b*, Sucrose gradient centrifugation of the α -latrotoxin receptor demonstrates comigration of the inhibitory activity with the 200K and 160K receptor polypeptides in the absence of α -latrotoxin (-LTx, upper lane) or bound to α -latrotoxin (+LTx, lower lane). Fractions were analysed for the α -latrotoxin receptor by immunoblotting (left) or for synaptotagmin-phosphorylation inhibitory activity (right).

METHODS. Synaptotagmin was purified as described in the legend to Fig. 2 except that CHAPS was used instead of Triton X-100. Purified synaptotagmin was dialysed against 0.9% CHAPS, 0.1 M Tris-HCl and 0.5 mM EDTA. About 0.1 μ g synaptotagmin was incubated in 80 mM Tris-HCl pH 7.4, 0.4 mM EDTA, 6 mM MgCl₂ and 4 mM CaCl₂ and 1 μ M γ [32 P]-ATP (specific activity 1.2 Ci mmol⁻¹) at 30 °C for 20 min. In the experiments shown in the left panel of *a*, about 0.3 μ g purified α -latrotoxin receptor (legend to Fig. 1) was added either before or after the incubations, and half of the incubations were cleaved with 40 ng trypsin at 37 °C for 10 min. In the experiments shown in the right panel of *a*, the indicated amounts of receptor were added before the phosphorylation reactions. Samples were analysed by SDS-PAGE followed by autoradiography (4 h exposure without screen), with the numbers on the right indicating positions of *M_r* markers ($\times 10^{-3}$). For sucrose gradient



fractionations, purified α -latrotoxin receptor was dialysed against 0.1 M Tris-HCl pH 7.4, 2 mM CaCl₂ and 0.9% CHAPS. Receptor was fractionated on a 5%-20% sucrose gradient (12 ml) in the same buffer with or without adding excess α -latrotoxin at 250,000g for 16 h. Fraction 1 presents the top of the gradient; in parallel runs *M_r* markers migrated at the following positions: Carbonic anhydrase (29K), fraction 2 and 3; bovine serum albumin (67K), fraction 4; alcohol dehydrogenase (150K), fraction 6; β -amylase (200K), fraction 8 and 9; apoferritin (443K), fractions 14 and 15. Fractions were analysed by immunoblotting using a α -latrotoxin receptor antibody (M.G. *et al.*, manuscript in preparation) or by phosphorylation inhibition of synaptotagmin as described above (exposure time 2 days without screen).

FIG. 4 Synaptotagmin phosphorylation in synaptic vesicles is selectively inhibited by the α -latrotoxin receptor. Synaptic vesicle proteins solubilized by CHAPS were phosphorylated with γ [32 P]-ATP *in vitro* in the presence or absence of Ca²⁺ and of α -latrotoxin receptor. After phosphorylation, the synaptotagmin from half of each sample was immunoprecipitated using a polyclonal antibody against synaptotagmin. Both the complete reactions and the immunoprecipitated synaptotagmin from each reaction were then analysed by SDS-PAGE and autoradiography, demonstrating selective inhibition of synaptotagmin phosphorylation by the α -latrotoxin receptor (right panel) but not of the phosphorylation of other synaptic vesicle proteins (left panel). **METHODS.** Synaptic vesicle proteins were solubilized in 2% (w/v) CHAPS, 20 mM Tris-HCl pH 7.4 and 1 mM EDTA on ice for 30 min, followed by centrifugation at 200,000g for 30 min. Phosphorylation reactions were performed for 20 min at 30 °C in a final volume of 12 μ l in a buffer containing in final concentrations: 80 mM Tris-HCl pH 7.4, 0.4 mM EDTA, 6 mM MgCl₂, 0.72% (w/v) CHAPS and 1 μ M γ [32 P]-ATP (specific activity: 2.8 Ci mmol⁻¹), with or without 3 mM Ca²⁺. Reactions contained 5.6 μ g synaptic vesicle protein and 0.2 μ g purified α -latrotoxin receptor were added as indicated. Phosphorylations were stopped by the addition of SDS to 1% (w/v) and boiling. Immunoprecipitations were performed using a polyclonal antibody against the C-terminal cytoplasmic domains of synaptotagmin¹². Films were exposed for 2 h (left panel) or 16 h (right panel) without screens.



We have demonstrated direct interaction between the α -latrotoxin receptor, a plasmamembrane protein that is probably exclusively localized to the presynaptic plasma membrane, and synaptotagmin, a synaptic vesicle membrane protein. No other synaptic protein interacted with the receptor, and the receptor seemed to interact with synaptotagmin through the cytoplasmic sequences of synaptotagmin. Synaptotagmin and the α -

latrotoxin receptor have complementary localizations in the nerve terminal, and both have been implicated in neurotransmitter release^{4,7}. Their interaction suggests a physiological role of the α -latrotoxin receptor as a docking protein for synaptic vesicle fusion through synaptotagmin. Furthermore, although α -latrotoxin in large doses has multiple nonspecific effects^{15,16}, low toxin concentrations cause rapid neurotransmitter release

by a mechanism that may involve the direct interaction of the cytoplasmic domains of the receptor with synaptic vesicles. Phosphorylation has been postulated to have a major role in regulating the efficiency of neurotransmitter release for example in long-term potentiation^{17,18}. With such a role, phosphorylation of synaptotagmin and its modulation by the α -latrotoxin receptor would be at a key point in the regulation of the efficiency of synaptic vesicle exocytosis. □

Received 30 May; accepted 19 July 1991.

- Clark, A. W., Mauro, A., Longenecker, H. E. Jr & Hurlbut, W. P. *Nature* **225**, 703–705 (1970).
- Frontali, N. et al. *J. Cell Biol.* **68**, 462–479 (1976).
- Pumplin, D. W. & Reese, T. S. *J. Physiol.* **273**, 443–457 (1977).
- Rosenthal, L. & Meldolesi, J. *Pharmac. Ther.* **42**, 115–134 (1989).
- Tzeng, M.-C. & Siekevitz, P. *J. Neurochem.* **33**, 263–274 (1979).
- Valtorta, F., Madeddu, L., Meldolesi, J. & Ceccarelli, B. *J. Cell Biol.* **99**, 124–132 (1984).
- Perin, M. S., Fried, V. A., Mignery, G. A., Jahn, R. & Südhof, T. C. *Nature* **345**, 260–261 (1990).
- Perin, M. S. et al. *J. Biol. Chem.* **266**, 615–622 (1991).
- Scheer, H. & Meldolesi, J. *EMBO J.* **4**, 323–327 (1985).
- Petrenko, A. G. et al. *EMBO J.* **9**, 2023–2027 (1990).
- Kikkawa, U., Kishimoto, A. & Nishizuka, Y. *A. Rev. Biochem.* **58**, 31–44 (1989).
- Perin, M. S., Brose, N., Jahn, R. & Südhof, T. C. *J. Biol. Chem.* **266**, 623–629 (1991).
- Südhof, T. C. & Jahn, R. *Neuron* **6**, 665–677 (1991).
- Wang, J. K. T., Walaas, S. I. & Greengard, P. *J. Neurosci.* **8**, 281–288 (1988).
- Okamoto, M., Longenecker, H. F. Jr & Riker, W. F. Jr *Science* **172**, 733–736 (1971).
- McMahon, H. T., Rosenthal, L., Meldolesi, J. & Nicholls, D. G. *J. Neurochem.* **55**, 2039–2047 (1990).
- Malinow, R. & Tsien, R. *Nature* **346**, 177–180 (1990).
- Bekkers, J. & Stevens, C. *Nature* **346**, 724–729 (1990).
- Kelly, R. B. *Neuron* **1**, 431–438 (1988).
- Knaus, P., Marquese-Pouey, B., Scherer, H. & Betz, H. *Neuron* **5**, 453–462 (1990).

ACKNOWLEDGEMENTS. This paper is dedicated to the late Y. A. Ovchinnikov who initiated part of this work. We thank J. L. Goldstein and M. S. Brown for critical advice, N. B. Levina for assistance in peptide sequencing, M. Cobb for helping with the phospho-amino acid analysis and K. Robinson, I. Leznicki and L. I. Pochtaryova for technical assistance.

CD4⁺ and CD8⁺ T cells acquire specific lymphokine secretion potentials during thymic maturation

Albert Bendelac & Ronald H. Schwartz

Laboratory of Cellular and Molecular Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 9000 Rockville Pike, Bethesda, Maryland 20892, USA

PERIPHERAL CD4⁺ and CD8⁺ T lymphocytes carry out different functions during immune reactions, partly as a result of the distinct patterns of lymphokines that they secrete upon stimulation. Using thymic cells from adult and newborn mice as well as from fetal organ cultures, we show here that this functional differentiation occurs inside the thymus and is completed during the single positive stage by the time the T-cell receptor becomes fully coupled to the intracellular activation pathways leading to lymphokine secretion. Surprisingly, CD4⁺ thymocytes differ from their immediate progeny, naive peripheral CD4⁺ cells, in that they secrete a broader range of lymphokines, including interleukins 4, 5 and 10 and γ -interferon, and more closely resemble immunologically experienced (activated or memory) CD4⁺ lymphocytes.

Upon primary stimulation, mature, resting peripheral CD8⁺ T cells in mice secrete large amounts of γ -interferon, a lymphokine involved in the defence against viruses, whereas CD4⁺ T cells mainly produce the lymphokine interleukin 2 (IL-2), an autocrine and paracrine T-cell growth factor. Preactivated or memory CD4⁺ T cells, however, can secrete a much larger array of lymphokines upon restimulation, including the lymphokines involved in class regulation of T and B cell responses: IL-4, IL-5, IL-10 and γ -interferon (γ -IFN)^{1–7}. To investigate the lineage relationships of the various functional subsets, we analysed the secretion potential of early T cells during thymic ontogeny.

CD4⁺8[−] and CD8⁺4[−] thymocytes from 6-week-old mice were purified and fractionated according to their surface expression of the heat stable antigen (HSA), a marker found at high levels on double-positive and early immature single-positive cells, and at low levels on mature thymocytes and peripheral T cells^{8,9}. Cell subsets were then stimulated, in the presence of accessory cells, with an anti- $\alpha\beta$ -T-cell receptor antibody coated on plastic microwells or, to bypass potential defects of T-cell receptor (TCR) coupling to the intracellular activation pathways in immature cells¹⁰, by a combination of the calcium ionophore ionomycin, and the phorbol ester phorbol myristate acetate (PMA). Figure 1 shows that immature (HSA^{high}) CD4⁺8[−] and CD8⁺4[−] thymocytes do not yet respond to signalling through their TCR by secreting lymphokines, but they both secrete large amounts of IL-2 as well as some γ -IFN in response to ionomycin and PMA. The two subsets diverge functionally at the mature (HSA^{low}) stage, the first stage at which TCR crosslinking induces abundant lymphokine production. We find that CD8⁺4[−] cells, like their peripheral counterparts, produce high levels of γ -IFN and low levels of IL-2. CD4⁺8[−] thymocytes secrete low levels of IL-2 and high levels of IL-4, IL-5, IL-10 and γ -IFN, a lymphokine pattern differing sharply from the one displayed by peripheral (mostly resting) CD4⁺ lymphocytes, which mainly produce high levels of IL-2. These dissimilarities were not related to differences in the kinetics of lymphokine secretion or to different dose-response curves for the stimulus used (data not shown). The same radical difference in patterns was observed in nine experiments where HSA^{low} CD4⁺8[−] thymocytes were directly compared with peripheral CD4⁺ lymphocytes from various lymphoid compartments, such as inguinal, axillary or mesenteric lymph nodes, spleen, blood or thoracic duct (Table 1). In fact, CD4⁺8[−] HSA^{low} thymocytes closely resemble the minor peripheral subpopulation of memory/preactivated CD4⁺ cells, as identified by its high levels of the phagocytic glycoprotein-1 (Pgp-1) surface marker and its low density. This subset accounts for almost all the production of lymphokines

TABLE 1 Functional similarity between mature HSA^{low} CD4⁺8[−] thymocytes and peripheral activated/memory CD4⁺ cells

Lymphokine	Production ratios	
	Thymic/peripheral CD4 ⁺ Mean of 9 experiments (range)	Large Pgp high/ small Pgp low peripheral CD4 ⁺
IL-2	0.17 (0.01–0.89)	0.46
γ -IFN	24 (4–192)	>35
IL-4	238 (32–677)	>152

The contrast between the lymphokine patterns elicitable from various T-cell populations is presented as the average ratio of the amounts of lymphokines produced upon stimulation with anti- $\alpha\beta$ TCR. Column 2: in nine experiments similar to that presented in Fig. 1, HSA^{low} CD4⁺8[−] thymocytes were compared with CD4⁺ lymphocytes from various lymphoid compartments (inguinal and axillary lymph nodes, $n=3$; mesenteric lymph nodes, $n=2$; spleen, $n=2$; blood, $n=1$; thoracic duct, $n=1$). Cells were stimulated with plate-bound anti- $\alpha\beta$ TCR antibody as described in Fig. 1, and the ratios of lymphokines secreted in the supernatants of thymic versus peripheral cell cultures was calculated to normalize for experiment-to-experiment variations. Results are presented as geometric means (ranges) of these ratios. The variations observed are in part accounted for by the variations in the proportions of activated or memory peripheral CD4⁺ lymphocytes. Column 3 compares activated/memory to resting peripheral CD4⁺ cells: peripheral CD4⁺ lymphocytes were obtained from spleens and fractionated on Percoll density gradients into high-density (Percoll 70%/66% interface) and low density (above Percoll 60% layer) cells, stained with anti-CD4-fluorescein isothiocyanate (FITC) and biotinylated anti-Pgp antibody (IM.78) followed by phycoerythrin-streptavidin. The CD4⁺ cells were then separated by fluorescence-activated cell sorting (FACS) according to their level of Pgp. Large Pgp^{high} cells produced 60 pM IL-2, 280 pM γ -IFN, and 24 pM IL-4. Small Pgp^{low} cells produced 130 pM IL-2, <8 pM γ -IFN and <0.16 pM IL-4.

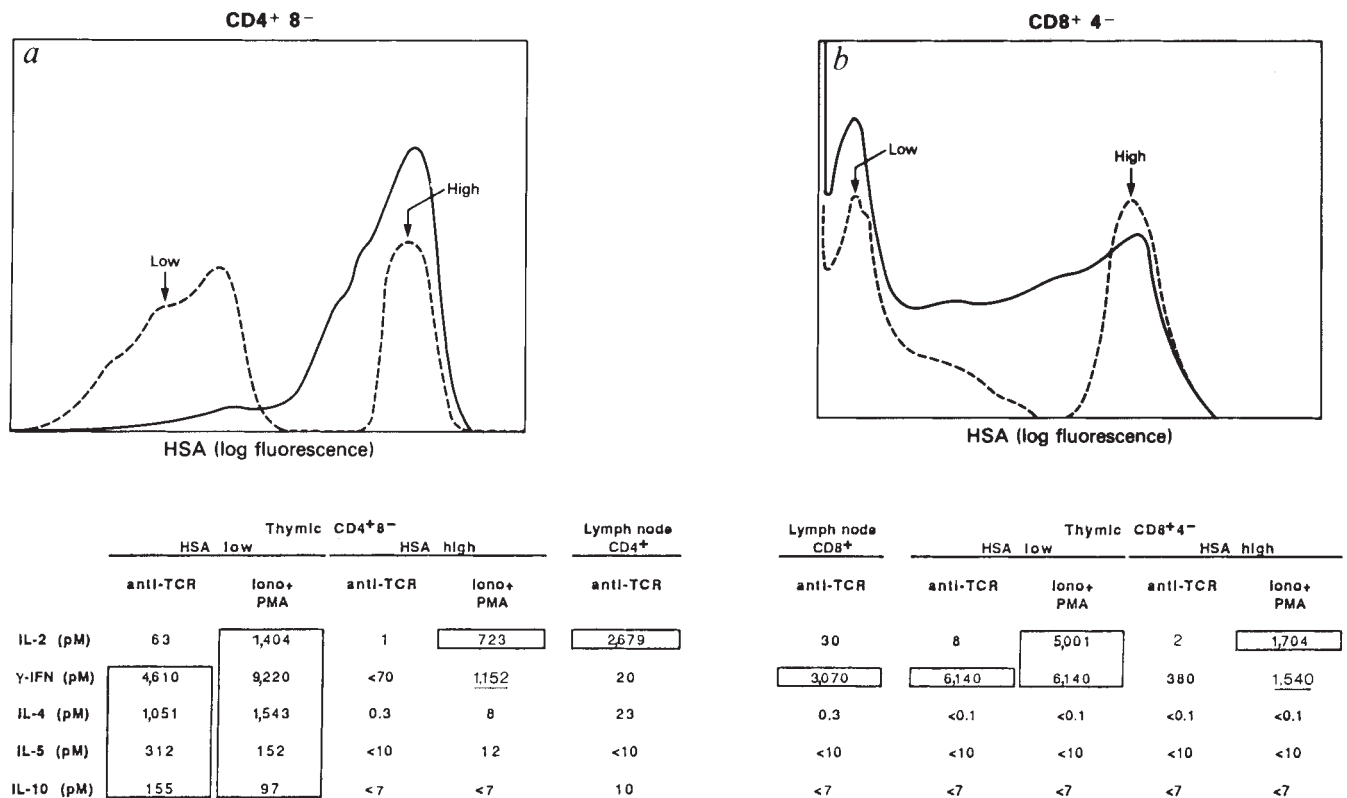


FIG. 1 Patterns of lymphokines secreted by HSA^{low} and HSA^{high} single-positive thymocytes. *a*, CD4⁺ 8⁻, and *b*, CD8⁺ 4⁻ thymocytes were stained for HSA (solid lines). HSA^{low} and HSA^{high} cells obtained after FACS-sorting (dashed lines) were stimulated by TCR crosslinking or the combination of ionophore and phorbol ester, and assayed for lymphokine secretion as shown in the table. The table combines two separate experiments: one involves the thymic CD4⁺ 8⁻ subset, and the other the thymic CD8⁺ 4⁻ subset along with the lymph node cells. To delineate general patterns of lymphokine secretion, the values within a 3- or 10-fold range of the maximum obtained in the experiment are respectively boxed or underlined.

METHODS. Thymuses from 6-week-old BALB/c mice were carefully dissected to remove perithymic tissues and thymocytes were passed twice on plastic dishes coated with anti-CD8 (2.43), or anti-CD4 (GK1.5) monoclonal antibodies; less than 1% of the recovered cells were double-positive. Cells were then stained with anti-CD8-FITC (53.6.7; Becton Dickinson), or anti-CD4-FITC (RM4-5; Pharmingen), and biotinylated anti-HSA (M1/69) followed by PE-conjugated streptavidin, and sorted (FACStar PLUS, Becton Dickinson) to obtain >98.5% pure single-positive cells expressing high or low levels of HSA. The proportion of HSA^{high} CD8⁺ 4⁻ cells expressing low or undetectable

levels of $\alpha\beta$ TCR was less than 22%. Peripheral CD4⁺ (99.8% pure) and CD8⁺ T cells (98.5% pure with 0.6% contaminating CD4⁺ cells) were sorted from mesenteric lymph node cells. Cells were plated at 1×10^5 in 0.25 ml medium with 2×10^4 irradiated (3,000 rads), dendritic cell-enriched accessory cells prepared from T-cell depleted spleen cells²¹. Stimulation was provided by anti- $\alpha\beta$ TCR antibody (H57; ref. 22) coated on plastic at $10 \mu\text{g ml}^{-1}$, or by the combination of $1 \mu\text{M}$ ionomycin (iono) and 1 ng ml^{-1} PMA. 48 h-culture supernatants were assayed for lymphokines using bioassays for IL-2 (CTLL line), IL-4 (CT.4S line; ref. 23), γ -IFN (protection of L929 cells from vesicular stomatitis virus; ref. 24) or sandwich enzyme-linked immunosorbent assays (ELISAs) for IL-5 and IL-10 (refs 25, 26). To ensure monospecificity, assays for IL-2 and IL-4 were routinely run in the presence of saturating amounts of anti-IL-4 (11B11; ref. 27) or a combination of anti-IL-2 and anti-IL-2 receptor (S4B6, PC61 and 7D4; refs 16, 28, 29) monoclonal antibodies, respectively. In selected cases, specificity was confirmed by blocking with the homologous antibody. The combination of ionomycin and PMA did not interfere with the bioassays at the concentrations used. Results are expressed in pM by comparison to recombinant standards run for each lymphokine.

other than IL-2 by peripheral CD4⁺ lymphocytes (Table 1, and refs 1-6). Ionomycin and PMA elicited roughly the same amounts of lymphokines as anti-TCR antibodies under optimal conditions, except for significantly increased levels of IL-2 (Fig. 1).

To investigate whether the mature phenotypes observed in adult thymuses are produced early in life, we purified the thymic single-positive subsets from newborn mice. Table 2 shows that the same lymphokine patterns were detectable from newborn cells after stimulation through their TCR. Because the level of IL-2 secreted by both subsets was limiting, owing to the small proportion of HSA^{low} mature cells in newborns, we also supplemented the cells with IL-2 to allow full expression of the potential lymphokine repertoire^{4,11}. Although IL-2 enhanced the secretion of those lymphokines already secreted by the thymic subsets, it did not modify the subset pattern: in particular, the addition of IL-2 enhanced by 64-fold the level of interferon produced by CD8⁺ 4⁻ cells but did not induce detectable IL-4 or IL-5. These results show that the neonatal and adult thymic

microenvironments equally support the differentiation of the lymphokine secretion potentials of CD4⁺ 8⁻ and CD8⁺ 4⁻ cells, even though they might differ, for example, for the development of T cells bearing the $\gamma\delta$ TCR (ref. 12).

Finally, as the possibility remained that the mature phenotypes in the thymus reflected the presence of a minor subpopulation of mature cells recirculating from the periphery¹³⁻¹⁵, we used an organ culture system to obtain functionally mature T cells *in vitro* from day-14 fetal thymuses. CD4⁺ 8⁻ and CD8⁺ 4⁻ cells were purified at the peak of their production, namely on days 22 and 29 respectively (data not shown), and tested for their pattern of lymphokine secretion upon TCR crosslinking. As shown in Fig. 2, the profiles were similar to those of newborns and adults. CD8⁺ 4⁻ cells mainly produced γ -IFN, whereas CD4⁺ 8⁻ cells secreted IL-4 in addition to γ -IFN. These results support the contention that no post-thymic maturation event is required for T cells to acquire their specific lymphokine potentials. A similar pattern was obtained in three additional experiments in which fetal organ

TABLE 2 Lymphokine production by newborn thymic single-positive cells

	Day-2 newborn thymic CD4 ⁺ 8 ⁻		Day-6 newborn thymic CD4 ⁺ 8 ⁻		Day-6 newborn thymic CD8 ⁺ 4 ⁻	
	anti-TCR	anti-TCR + IL-2	anti-TCR	anti-TCR	anti-TCR + IL-2	anti-TCR + IL-2
IL-2 (pM)	1	ND	0.7	0.1	ND	ND
γ -IFN (pM)	<24	192	384	384	24,576	24,576
IL-4 (pM)	7	34	49	<0.1	<0.1	<0.1
IL-5 (pM)	24	ND	500	<12	<12	<12

Two-day-old BALB/c and six-day-old (BALB/c $\times$ B6) thymocytes were processed as described in Fig. 1 and stained with anti- $\gamma\delta$ TCR-phycoerythrin (PE) (GL3, Pharmingen) and either anti-CD4-FITC plus anti-CD8-PE or anti-CD8-FITC plus anti-CD4-PE respectively. FITC positive-PE negative cells were sorted to obtain the single-positive subsets, and assayed for lymphokine production after stimulation with anti- $\alpha\beta$ TCR in the presence of accessory cells as described in Fig. 1. Six-day-old thymic CD4⁺8⁻ enriched cells were obtained by removing CD8⁺4⁻ cells and double-positive cells by two rounds of panning on anti-CD8-coated plates. Exogenous human recombinant IL-2 was from Hoffman La Roche (Basel, Switzerland) added at a final concentration of 10 units per ml (30 pM). ND, not determined.

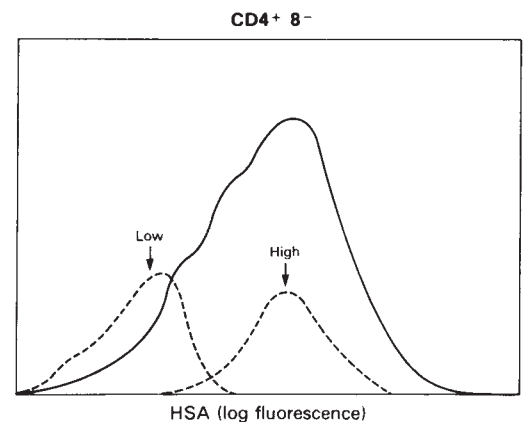
culture derived HSA^{low} CD4⁺8⁻ thymocytes were tested, including one experiment in which IL-5 was measured and found also to be produced. Interestingly, when the organ culture results were compared with fresh adult peripheral or thymic cells, we consistently found that the relative amounts of IL-2, IL-4 and γ -IFN produced by organ-culture derived HSA^{low} CD4⁺8⁻ cells were in between the amounts produced by these two populations. This observation may indicate that the activated phenotype of HSA^{low} CD4⁺8⁻ thymocytes is only transient and that it spontaneously reverts to the resting, IL-2-only producing phenotype normally observed after export to the periphery, a transition that can be captured during organ culture. Alternatively, it is possible that distinct functional lineages are produced in the thymus. In particular, if the precursors to most of the peripheral resting CD4⁺ cells (high IL-2 producers) are normally exported out of the thymus faster than a minor subset of differentiated lymphokine producers, the latter might quantitatively predominate inside the thymus. Kinetic studies and limiting dilution analyses are in progress to resolve this issue.

Our results establish that the different patterns of lymphokines characteristic of peripheral CD4⁺ and CD8⁺ cells are acquired

during progression from immature HSA^{high} to mature HSA^{low} thymocytes by the time the TCR becomes functionally linked to the intracellular activation pathways leading to lymphokine production. The secretion of the full range of the five tested lymphokines by CD4⁺8⁻ thymocytes was an unexpected result, because only IL-2 is produced at high levels by peripheral resting T cells, and the production of high levels of IL-4, IL-5, IL-10 and γ -IFN has only been obtained from activated or memory cells *in vivo*¹⁻⁷, or by cultured T-cell lines^{16,17}. It is conceivable that the thymic CD4⁺8⁻ phenotype might represent a distinct functional lineage, responsible for the production of lymphokines other than IL-2 during immune responses. Alternatively, it might result from a form of tolerance induction (immune deviation), imparted to a subset of self-reactive cells that have escaped deletion. We favour the hypothesis, however, that the functionally activated phenotype of the CD4⁺8⁻ subset, like its expression of the CD45 isoforms also found on activated but not resting peripheral lymphocytes (ref. 18, and data not shown), reflects a previous stimulation through the TCR during a positive selection step on major histocompatibility complex molecules¹⁹. Following migration to the periphery, the thy-

FIG. 2 Lymphokine secretion potential of single-positive cells from fetal organ culture. The top panel shows the distribution of membrane HSA among CD4⁺8⁻ cells obtained from organ-cultured fetal thymuses before (solid lines) and after (dashed lines) sorting into HSA^{low} and HSA^{high} fractions. The lymphokine profiles, obtained after stimulation by TCR crosslinking, are shown in the table for organ culture-derived cells and fresh adult lymphocytes.

METHODS. 14-day-old fetal thymuses from timed-pregnant C57Bl/6 mice were cultured on 24-mm diameter microporous membranes (Transwell, Costar) in contact with Iscove's medium supplemented with 10% fetal calf serum, glutamine, 5×10^{-5} M 2-mercaptoethanol and antibiotics in a 5% CO₂ atmosphere. CD4⁺8⁻ cells were obtained after 8 days in culture by panning on anti-CD8 coated plates, staining with anti-CD4-FITC and anti-CD8-PE and sorting. For fractionation according to the level of membrane HSA, the cells were additionally stained with biotinylated anti-HSA and streptavidin-allophycocyanin (Molecular Probes). CD8⁺4⁻ cells were sorted after 15 days of organ culture. Adult CD4⁺ populations included anti-CD8 (3.155) plus anti-HSA (J11d.2) plus complement-treated thymic cells or anti-CD8 plus complement-treated and nylon-wool passed mesenteric lymph node cells from C57Bl/6 mice. Cells were stimulated for 48 h in wells coated with anti- $\alpha\beta$ TCR antibody as described in the legend to Fig. 1 at 1×10^5 cells per well, except for the minor CD4⁺8⁻ HSA^{low} subset assayed at 0.33×10^5 cells per well and



	Adult		Fetal thymic organ culture			
	Lymph node CD4 ⁺	Thymic CD4 ⁺ 8 ⁻ HSA low	CD4 ⁺ 8 ⁻ HSA low	DAY 22 CD4 ⁺ 8 ⁻ HSA high	CD4 ⁺ 8 ⁻ unseparated	DAY 29 CD8 ⁺ 4 ⁻ unseparated
IL-2 (pM)	149	2	256	4	18	50
γ -IFN (pM)	82	776	125	16	39	15432
IL-4 (pM)	2	773	70	0.6	5	<0.05

therefore adjusted by a 3-fold multiplication. Lymphokines were assayed as described in Fig. 1, except for γ -IFN, which was measured with a sandwich ELISA³⁰.

mocytes could follow two paths: in the absence of antigen they would revert to the resting phenotype, but in the presence of antigen they would become immediately engaged in an immune response, where they could be the elusive^{11,20} primary source of IL-4. □

Received 11 March; accepted 12 July 1991.

- Budd, R. C., Cerottini, J.-C. & MacDonald, H. R. *J. Immunol.* **138**, 3583–3586 (1987).
- Swain, S. L., McKenzie, D. T., Weinberg, A. D. & Hancock, W. *J. Immunol.* **141**, 3445–3455 (1988).
- Hayakawa, K. & Hardy, R. R. *J. exp. Med.* **168**, 1825–1838 (1988).
- Ben-Sasson, S. Z., Le Gros, G., Conrad, D. H., Finkelman, F. D. & Paul, W. E. *J. Immunol.* **145**, 1127–1136 (1990).
- Lee, W. T., Yin, X.-M. & Vitetta, E. S. *J. Immunol.* **144**, 3288–3295 (1990).
- Ernst, D. N. et al. *J. Immunol.* **145**, 1295–1302 (1990).
- Bottomly, K. et al. *Eur. J. Immunol.* **19**, 617–623 (1989).
- Crispe, I. N. & Bevan, M. J. *J. Immunol.* **138**, 2013–2018 (1987).
- Ceredig, R. *J. Immunol.* **141**, 355–362 (1988).
- Finkel, T. H., McDuffie, M., Kappler, J. W., Marrack, P. & Cambier, J. C. *Nature* **330**, 179–181 (1987).
- Le Gros, G., Ben-Sasson, S. Z., Seder, R., Finkelman, F. D. & Paul, W. E. *J. exp. Med.* **172**, 921–929 (1990).
- Ikuta, K. et al. *Cell* **62**, 863–874 (1990).
- Michie, S. A., Kirkpatrick, E. A. & Rouse, R. V. *J. exp. Med.* **168**, 1929–1934 (1988).
- Grossmann, J., Lohler, J. & Lehmann-Grube, F. *J. Immunol.* **146**, 293–297 (1991).
- Agus, D., Surh, C. D. & Sprent, J. *J. exp. Med.* **173**, 1039–1046 (1991).
- Mosmann, T. R., Chervinski, H., Bond, M. W., Giedlin, M. A. & Coffman, R. L. *J. Immunol.* **136**, 2348–2357 (1986).
- Fiorentino, D., Bond, M. W. & Mosmann, T. R. *J. exp. Med.* **170**, 2081–2095 (1989).
- Pilarski, L. M., Gillitzer, R., Zola, H., Shortman, K. & Scollay, R. *Eur. J. Immunol.* **19**, 589–597 (1989).
- Bevan, M. J. & Fink, P. *J. Immunol. Rev.* **42**, 3–26 (1978).
- Swain, S. L., Weinberg, A. D. & English, M. J. *Immunol.* **144**, 1788–1799 (1990).
- Nussenzweig, M. C. & Steinman, R. M. *J. exp. Med.* **151**, 1196–1212 (1980).
- Kubo, R. T., Born, W., Kappler, J. W., Marrack, P. & Pigeon, M. *J. Immunol.* **142**, 2736–2742 (1989).
- Hu, L., Ohara, J., Watson, C., Tsang, W. & Paul, W. E. *J. Immunol.* **142**, 800–807 (1989).
- Dahl, H. & Degre, M. *Acta path. microbiol. Scand.* **80**, 863–870 (1972).
- Schumacher, J. H. et al. *J. Immunol.* **141**, 1575–1581 (1988).
- Mosmann, T. R. et al. *J. Immunol.* **145**, 2938–2945 (1990).
- Ohara, J. & Paul, W. E. *Nature* **315**, 333–336 (1985).
- Ceredig, R., Lowenthal, J. W., Nabholz, M. & MacDonald, H. R. *Nature* **314**, 98–100 (1985).
- Ortega, G. R., Robb, R. J., Shevach, E. M. & Malek, T. R. *J. Immunol.* **133**, 1970–1982 (1984).
- Chervinski, H. M., Schumacher, J. H., Brown, K. D. & Mosmann, T. R. *J. exp. Med.* **166**, 1229–1244 (1987).

ACKNOWLEDGEMENTS. We thank P. Matzinger and W. E. Paul for discussion; R. Asofsky, E. J. Fuchs, R. N. Germain, P. Matzinger, W. E. Paul and S. S. Stringfellow for reviewing the manuscript; K. W. Moore, R. L. Coffman and W. E. Paul for making available to us the reagents for the IL-10 and IL-5 ELISAs and the CT-4S cell line, respectively; C. Chen for technical help; J. Hooley for electronic cell sorting, and P. Matzinger for thoracic duct cannulation.

Bcl-2 maintains B cell memory

Gabriel Nuñez, David Hockenbery,
Timothy J. McDonnell, Craig M. Sorensen
& Stanley J. Korsmeyer

Howard Hughes Medical Institute and the Departments of Medicine,
Molecular Microbiology and Pathology, Washington University School of
Medicine, Saint Louis, Missouri 63110, USA

THE number of lymphocytes in an animal is remarkably constant despite antigen-driven proliferation and a high rate of B-cell lymphopoiesis. This reflects the relatively brief lifespan of many newly generated B cells and argues for a well-regulated death mechanism^{1–3}. Even so, a secondary immune response can be generated years after a primary exposure to antigen⁴. Antigen that might restimulate B cells persists for extended periods on follicular dendritic cells in the light zone of germinal centres^{5–13}. Antigen-binding B cells have also been found months after the end of obvious cell division¹⁴. The precise signal that enables certain B cells to emerge as long-term surviving memory cells^{14–17} is unknown. Bcl-2, an inner mitochondrial membrane protein¹⁸, blocks programmed cell death in B cells^{18–20}. We report here that this proto-oncogene maintains immune responsiveness. Transgenic mice overproducing Bcl-2 have a long-term persistence of immunoglobulin-secreting cells and an extended lifetime for memory B cells.

A transgenic mouse bearing a Bcl-2-Ig minigene overproduces Bcl-2 (refs 21, 22) and provided a model to assess its role in the intact immune system. Transgenics developed a strong

primary response to a T-dependent antigen, FITC-KLH (fluorescein isothiocyanate-keyhole limpet haemocyanin), when assessed at day 7 after immunization. Six transgenics averaged 426 direct (IgM) plaque-forming cells (p.f.c.) per 10⁶ spleen cells and 476 indirect (IgM+IgG) p.f.c. per 10⁶ spleen cells. Control littermates displayed a comparable amount of total, indirect p.f.c. at 351 per 10⁶ cells, whereas their direct p.f.c. were 190 per 10⁶ cells. The predominance of IgM in the primary response by transgenics is consistent with their expanded IgM/IgD B cells²². Moreover, the capacity of transgenics to mount a classic secondary response 4 weeks after the primary immunization was normal (Fig. 1). Both indirect p.f.c.s (259 versus 206) and direct p.f.c.s (59 versus 67) were comparable for controls and transgenics, respectively. Most of these plaques were indirect, indicating that the immunoglobulin-secreting cells (ISC) had undergone a heavy-chain class switch to IgG.

We wished to assess whether overexpressed Bcl-2 altered the kinetics of antigen responsiveness. Normal animals demonstrated a sequential decline in ISC over time (Fig. 1). By contrast, transgenics possess fivefold (216 versus 41) at day 25 and 20-fold (349 versus 15) at day 75 more p.f.c. than control littermates (Fig. 1). Parallel increases in levels of specific antibody were found in serum of transgenics when assessed by enzyme-linked immunosorbent assay (ELISA) (data not shown). Thus, deregulated Bcl-2 extends the lifetime of an ISC population.

To assess whether Bcl-2 could influence the generation and maintenance of memory B cells, an adoptive transfer²³ protocol was designed that transferred primed B cells to an antigen-naïve recipient (Fig. 2). Donor mice were immunized with FITC-KLH to generate memory B cells. A Percoll density gradient enriched for resting memory B cells and minimized the transfer of residual antigen. Recipients had been immunized with an unrelated carrier protein, bovine serum albumin (BSA), to generate excess

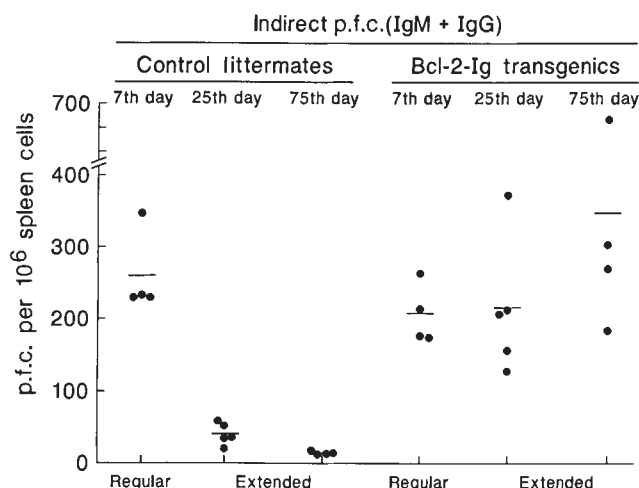


FIG. 1 Secondary immune response to FITC-KLH by Bcl-2-Ig and control littermate mice. Mice were immunized with 50 µg of FITC-KLH and a secondary challenge was administered 4 weeks later. Specific anti-FITC p.f.c. were counted at 7, 25 and 75 days after the secondary antigen challenge. METHODS. FITC was coupled to sheep erythrocytes (SRBC) (Colorado Serum, Denver). Spleen cells were incubated with FITC-SRBC in agar at 37 °C for 60 min and plated at 1×10^6 cells per slide. Indirect (IgM+IgG) plaques were developed by adding a facilitating rabbit anti-mouse Ig serum (a gift of C. Pierce, Washington University) for 90 min at 37 °C. The incubation was prolonged for another 90 min at 37 °C in the presence of guinea-pig complement (Rockland, Gilbertville, Pennsylvania). All p.f.c. assays were performed in duplicate. FITC-specific p.f.c. were determined from the mean of duplicate slides after subtraction of background p.f.c. Background represented the plaques formed in the presence of excess free antigen, 25 µg ml⁻¹ FITC-coupled to ε-amino-n-caproic acid. Male and female Bcl-2-Ig (5–10-weeks old) and control littermates were injected intraperitoneally with 50 µg FITC-KLH as an Alum precipitate. Secondary challenge with FITC-KLH was done in the same manner 4 weeks after the primary immunization.

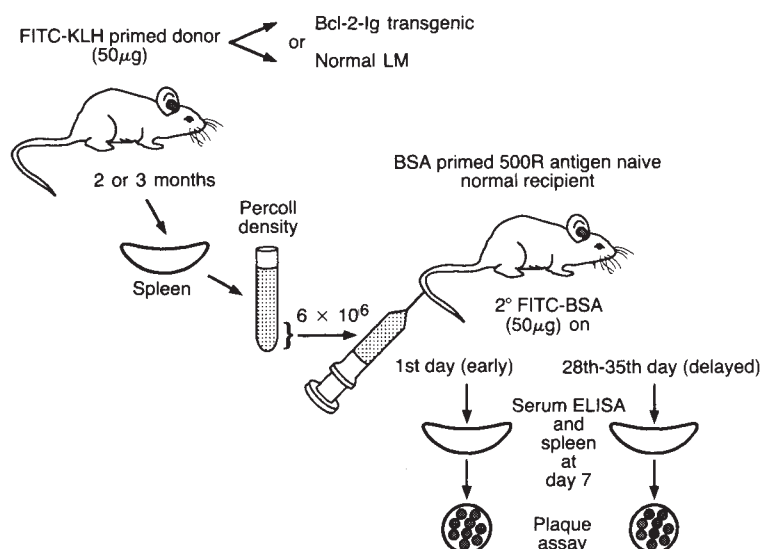


FIG. 2 Adoptive transfer protocol.

helper T cells. Sublethal irradiation enabled transferred B cells to populate the recipient spleen. Recipients receive the secondary challenge with FITC-BSA. The substituted carrier, BSA, ensured that transferred B and not T memory would be assessed.

Control experiments demonstrated that only transferred B cell memory was being analysed (Fig. 3a). Recipients of primed B cells from either normal littermates or Bcl-2-Ig transgenics generated equivalent antibody responses when challenged with antigen at the time of transfer. Memory cells from transgenics generated 433 indirect p.f.c. per 10^6 spleen cells, whereas cells from normal littermates resulted in 396 indirect p.f.c. per 10^6 cells (Fig. 3a). Most p.f.c. derived from transferred memory cells were secreting IgG. But when the secondary immunization was delayed for 28–35 days, recipients of memory cells from normal mice fell precipitously to minimal responses of 20 p.f.c. per 10^6 spleen cells (Fig. 3a). This confirms that the half-life of transferred B memory cells in the absence of antigen is remark-

ably short^{24,25}. By contrast, memory cells from Bcl-2-Ig mice persisted and still generated 254 indirect p.f.c. per 10^6 and 71 direct p.f.c. per 10^6 spleen cells when restimulated with FITC-BSA 4–5 weeks after transfer. Serum assays confirmed this difference with transgenic transfers demonstrating high titre IgG-specific antibody (1/2,400), whereas recipients of normal cells had declined to a residual IgG antibody titre of 1/420 (Fig. 3b).

Little is known about the genetic events that enable memory lymphocytes to exist over long periods. Bcl-2-Ig mice had relatively normal primary and initial secondary responses during a period when antigen is not limiting. Yet deregulated Bcl-2 overrode the need for persistent antigen in the maintenance of B-cell memory in the setting of an adoptive transfer. This suggests that the B-cell survival pathway may involve antigen-induced reactivation of Bcl-2. Bcl-2 protein is topographically restricted in germinal centres to the follicular mantle²⁶, and to portions of

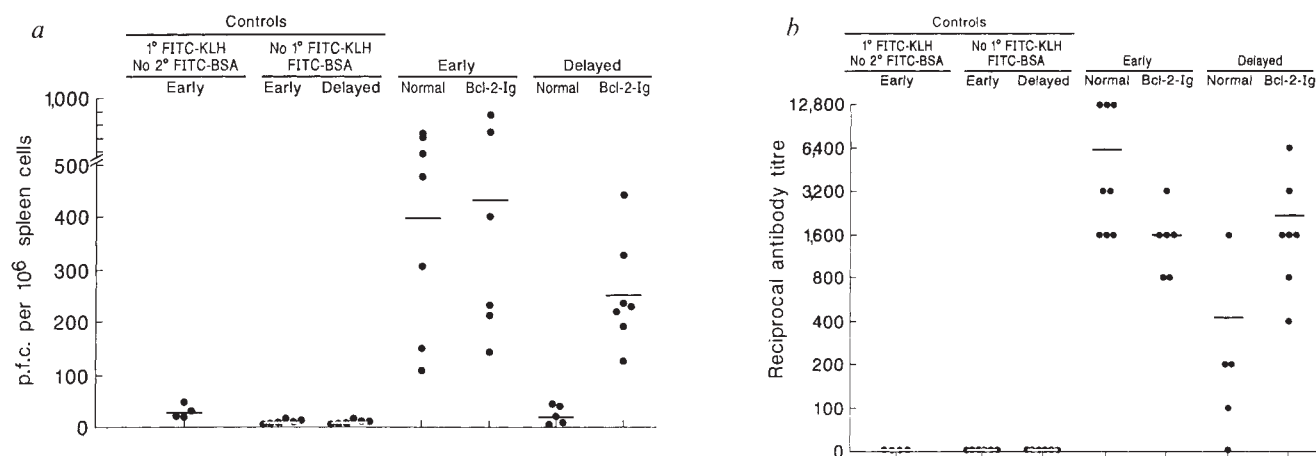


FIG. 3 a, B-cell memory response to FITC. Controls indicate that primed donor cells from transgenic mice when transferred to recipients which were not given the secondary antigen, resulted in no substantial p.f.c. (1st column). This indicates that resting, nonsecreting B cells rather than ISC were being transferred. Conversely, nonimmune B cells were transferred to recipients that received FITC-BSA. These recipients developed no p.f.c. response indicating that sublethal irradiation had eliminated the recipient's endogenous B cell response (2nd column). Adoptive host mice that received primed lymphocytes from either normal or Bcl-2-Ig mice were assessed for FITC-specific p.f.c. The secondary immunization was administered to the recipients on the day of transfer (early) (3rd column) or 28–35 days later (delayed) (4th column). b, Adoptive host mice that received primed lymphocytes from either normal or Bcl-2-Ig mice. The serum concentration of IgG-specific antibody was determined by ELISA 7 days after the secondary challenge with FITC-BSA.

METHODS. Donor mice were assessed for b/k heterozygosity at the *H-2* region by

DNA typing. Spleen cells from FITC-KLH primed or nonimmunized donors were first enriched for small resting B cells on a Percoll gradient by collecting fractions with a density $>1.085 \text{ g ml}^{-1}$ ²². Spleen cell preparations from Bcl-2-Ig and control mice were $56 \pm 4\%$ and $43 \pm 4\%$ positive, respectively, with the rat monoclonal antibody 187.1 specific for mouse κ light chain. Cells (6×10^6) were injected intravenously into the tail veins of 10–15-week-old C57BL6 $\times$ C3H (b/k) male recipients that had been primed with 100 μg BSA in complete Freund's adjuvant. Recipient mice had been irradiated (500 rads) by a Gamma Cell 40 Cs source (Atomic Energy of Canada) 24 h before cell transfer to enable transferred B cells to populate the spleen. Recipient mice received the secondary challenge with 50 μg of FITC-BSA as alum precipitate given intraperitoneally either on the day of transfer (early) or 28–35 days later (delayed). Seven days after FITC-BSA administration spleens were removed and serum samples obtained.

the light zone²⁷ implicated in the selection and maintenance of plasma cells and memory B cells. Treatment of centrocytes with anti-immunoglobulin plus anti-CD40 rescues them from apoptosis¹². Moreover, crosslinking surface immunoglobulins plus B-cell growth factor increases Bcl-2 RNA²⁸. The effects of Bcl-2 on immune responses support a role for Bcl-2 in long-lived lymphocytes responsible for antibody production and memory. □

Received 5 March; accepted 9 July 1991.

- Opstelt, D. & Osmond, D. G. *J. Immunol.* **131**, 2635–2640 (1983).
- Gray, D. *J. exp. Med.* **167**, 805–816 (1988).
- Forster, I. & Rajewsky, K. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4781–4784 (1990).
- Berek, C., Jarvis, J. M. & Milstein, C. *Eur. J. Immunol.* **17**, 1121–1129 (1987).
- Tew, T. G. & Mandel, T. E. *Immunology* **37**, 69–77 (1979).
- Klaus, G. G. B., Humphrey, J. G., Kunkle, A. & Dougworth, D. W. *Immunol. Rev.* **53**, 3–28 (1980).
- Coico, R. F., Bhogel, B. S. & Thorbecke, G. J. *J. Immunol.* **131**, 2254–2257 (1983).
- Kraal, G., Weissman, I. L., Butcher, L. C. *Cell. Immunol.* **115**, 78–87 (1988).
- Kosko, M. H., Burton, G. F., Kapasi, Z. F., Szakal, A. K. & Tew, J. G. *Immunology* **68**, 312–318 (1989).
- Kraal, G., Weissman, I. L. & Butcher, E. C. *Nature* **298**, 377–379 (1982).
- MacLennan, I. C. M. & Gray, D. *Immunol. Rev.* **91**, 61–85 (1986).
- Liu, Y.-J. *et al. Nature* **342**, 929–931 (1989).
- Fliedner, T. M., Kesse, M., Cronkite, E. P. & Robertson, T. S. *Ann. N.Y. Acad. Sci.* **13**, 578–594 (1964).
- Schittke, B. & Rajewsky, K. *Nature* **346**, 749–751 (1990).
- Gowans, J. L. & Uhr, J. W. *J. exp. Med.* **124**, 1017–1030 (1966).
- Colle, J. H., Truffe-Bachi, P. & Freitas, A. *Eur. J. Immunol.* **18**, 1307–1314 (1988).
- Linton, P. J., Decker, D. J. & Kliman, N. R. *Cell* **59**, 1049–1059 (1989).
- Hockenbery, D., Nunez, G., Millman, C., Schreiber, R. D. & Korsmeyer, S. J. *Nature* **348**, 334–338 (1990).
- Vaux, D. L., Cory, S. & Adams, J. M. *Nature* **335**, 440–442 (1988).
- Nunez, G. *et al. J. Immunol.* **144**, 3602–3610 (1990).
- McDonnell, T. J. *et al. Cell* **57**, 79–88 (1989).
- McDonnell, T. J. *et al. Molec. cell. Biol.* **10**, 1901–1907 (1990).
- Siekevitz, M., Kocks, C., Rajewsky, K. & Dildrop, R. *Cell* **48**, 757–770 (1987).
- Celada, F. *J. exp. Med.* **125**, 199–212 (1967).
- Gray, D. & Skarvall, H. *Nature* **336**, 70–72 (1988).
- Pezella, F. *et al. Am. J. Path.* **137**, 225–232 (1990).
- Hockenbery, D. M., Zutter, M., Hickey, W., Naim, M. & Korsmeyer, S. J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Grüniger, W. G., Seto, M., Boutain, B., Goldman, P. & Korsmeyer, S. J. *J. clin. Invest.* **80**, 1512–1515 (1987).

ACKNOWLEDGEMENTS. We thank C. Davinroy for preparation of this manuscript and C. Pierce and J. Kapp for helpful discussions and reagents. This work was supported by the American Cancer Society (D.H.) and the NIH.

Direct access to serum macromolecules by intraerythrocytic malaria parasites

B. Pouvelle*, R. Spiegel*, L. Hsiao*, R. J. Howard†, R. L. Morris*, A. P. Thomas* & T. F. Taraschi*‡

* Department of Pathology and Cell Biology, Jefferson Medical College, 1020 Locust Street, Philadelphia, Pennsylvania 19107, USA
† DNAX Research Institute, Palo Alto, California 94304, USA

TRAFFICKING pathways in malaria-infected erythrocytes are complex because the internal parasite is separated from the serum by the erythrocyte and parasitophorous vacuolar membranes¹. Intraerythrocytic *Plasmodium falciparum* parasites can endocytose dextrans, protein A and an IgG_{2a} antibody. Here we show that these macromolecules do not cross the erythrocyte or parasitophorous vacuolar membranes, but rather gain direct access to the aqueous space surrounding the parasite through a parasitophorous duct. Evidence for this structure includes visualization of membranes that are continuous between the parasitophorous vacuolar and erythrocyte membranes, and surface labelling of the parasite with fluorescent macromolecules under conditions that block endocytosis. The parasite can internalize by fluid-phase endocytosis macromolecules from the aqueous compartment surrounding it. Thus, surface antigens on trophozoites and

schizonts should be considered as targets for antibody-directed parasitocidal agents.

We have shown that fluorescent phosphatidylcholine added to tissue culture medium containing trophozoite-stage *P. falciparum*-infected erythrocytes rapidly labels parasite membranes under conditions that block endocytosis^{2,3}. To determine how the parasites are labelled, we compared the uptake of large fluid-phase markers into infected red blood cells under conditions that block (4 °C, ATP depletion)⁴ or permit endocytosis.

Cell suspensions of trophozoite-stage *P. falciparum*-infected cells were incubated with rhodamine-dextran (relative molecular mass, 10,000 (10K)) at 37 °C and washed. Confocal fluorescence imaging microscopy revealed heavy labelling of the parasite after 5 min incubation (Fig. 1a). When infected cells were incubated with rhodamine-dextran at 4 °C and washed, no cell-associated fluorescence was detected (Fig. 1b). Similar labelling patterns were obtained with a 70K rhodamine-dextran. This result could be due either to the blocking of fluid-phase endocytosis at reduced temperature or to the fact that at 4 °C the uncharged rhodamine-dextran enters the erythrocyte but cannot bind to the membrane and is removed by washing. To distinguish between these possibilities, we compared the labelling at 37 °C and 4 °C with rhodamine- and fluorescein-dextrans (10 or 70K) that have several positively charged binding sites (lysine residues) for binding to negatively charged membranes. Figure 1c and d show labelling of the parasite under both conditions. The parasite is labelled more heavily at 37 °C than at 4 °C. We obtained similar results with the large (42K) globular protein rhodamine-protein A. Confocal fluorescence images of infected cells incubated in tissue culture with fluorescein-protein A for 2 h at 37 °C and 4 °C are shown in Fig. 2a and b, respectively: the parasites are labelled at both temperatures.

To distinguish external and internal membrane labelling, we incubated cells labelled with fluorescein-protein A or fluorescein-dextran-lysine with an anti-fluorescein IgG at 4 °C. When this antibody binds fluorescein, more than 95% of the fluorescence is quenched. It is apparent from Fig. 2c that parasites labelled at 37 °C with fluorescein-protein A are fluorescent, whereas those labelled at 4 °C are not (Fig. 2d). Similar results

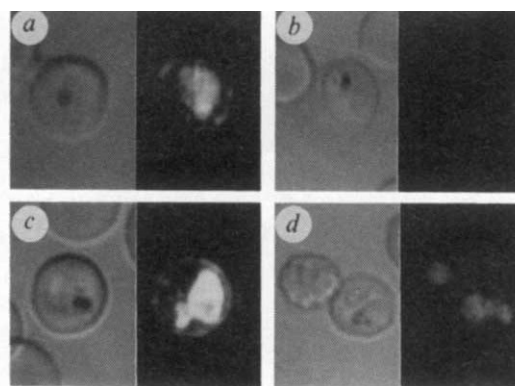


FIG. 1 Confocal fluorescence and transmitted light images of malaria-infected red blood cells incubated in tissue culture with a, rhodamine-dextran (10K) for 2 h at 37 °C, b, rhodamine-dextran (10K) for 2 h at 4 °C; c, rhodamine-dextran-lysine (10K) for 2 h at 37 °C; and d, rhodamine-dextran-lysine (10K) for 2 h at 4 °C. In all cases, cells were washed 5 times at 4 °C before microscopy. The C-5 clone of the FCR-3/FMG African strain of *P. falciparum*¹⁵ was cultured *in vitro* as described¹⁶. Ring-stage Gambian human isolates 249, 251, K86 and K41 were cultured by this method for only one cycle. Labelling patterns in the continuously cultured parasites and the human isolates were similar. Uninfected erythrocytes did not take up dextran, as mature erythrocytes do not endocytose. Cell viability was always greater than 98% using a live/dead viability kit (Molecular Probes, Oregon). Fluorescence images were obtained with a Bio-Rad MRC 500 confocal imaging system using an Olympus IMT-2 inverted microscope.

‡ To whom correspondence should be addressed.

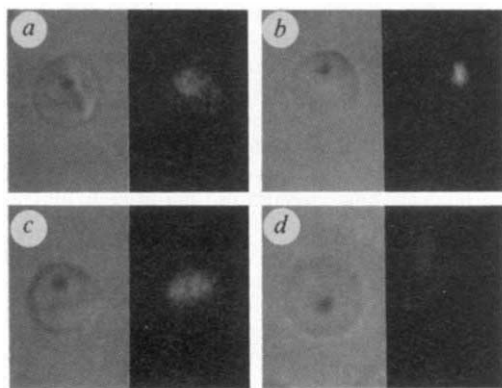


FIG. 2 Transmitted light and confocal fluorescence images of malaria-infected red blood cells incubated in tissue culture with fluorescein-protein A (10 mg ml^{-1} ; Sigma) for 2 h at *a*, 37°C or *b*, 4°C . Cells were washed 5 times at 4°C before microscopy. Following confocal fluorescence imaging microscopy, 1 U of an anti-fluorescein IgG antibody was added to *c*, cells from *a*, and *d*, cells from *b* at 4°C . To ensure that labelling of the parasite at 4°C was not caused by an alteration of the structure or permeability properties of the erythrocyte by the fluorescent dye bound to protein A, erythrocytes were incubated for 2 h at 4°C with unlabelled protein A-biotin. Cells were washed 4 times and incubated with rhodamine-avidin at 4°C . The labelling patterns (not shown) were identical to those obtained when fluorescein-protein A was the primary label; the parasitophorous vacuolar membrane and parasite plasma membrane were labelled at 4°C .

were obtained for the fluorescein-dextran-lysine-labelled cells. These results demonstrate that at 4°C , only the external membranes (erythrocyte, parasitophorous vacuolar and parasite plasma membranes) were labelled, whereas at 37°C internal parasite structures were also labelled.

To identify membranes within infected red blood cells, living, trophozoite-stage infections were labelled with either C_6 -NBD-PC (1-acyl-2-(*N*-4-nitrobenzo-2-oxa-1,3-diazole)-aminocaproyl phosphatidylcholine) (Fig. 3*a,b*), C_6 -NBD-phosphatidylethanolamine (PE) (Fig. 3*c*) or C_6 -NBD-ceramide (Fig. 3*d*) and examined by confocal fluorescence imaging microscopy (CFIM). Irrespective of the phospholipid headgroup, tubular extensions of membrane appeared to bridge the erythrocyte and parasite vacuolar membranes. These structures, here called 'parasitophorous ducts', could interconnect the parasitophorous vacuole and the external medium to provide a path by which macromolecules in the serum can passively diffuse to the parasite. Using NBD-PC, Grellier *et al.*⁵ have observed tubular structures protruding from the parasitophorous vacuolar membrane, which they suggest might provide a transient pathway

for the inward exchange of serum phospholipids from the erythrocyte membrane to the parasite at points of membrane contact.

Visualization of the parasitophorous duct is not restricted to NBD-labelled phospholipids. When infected erythrocytes are incubated at 4°C or 37°C with 30-nm, highly charged fluorescent latex beads, tubular extensions from the parasitophorous vacuole are seen to traverse the erythrocyte cytoplasm (Fig. 3*e,f*).

Several lines of evidence suggest that once serum macromolecules diffuse through the parasitophorous duct, they can be endocytosed by the parasite at 37°C . Parasites in infected erythrocytes incubated with rhodamine-dextran at 37°C are fluorescent, whereas those incubated at 4°C are not. Parasites in infected red blood cells incubated with rhodamine-dextran-lysine, rhodamine-protein A, or fluorescein-protein A are more heavily labelled at 37°C than at 4°C . Furthermore, trophozoite-stage parasites incubated *in vitro* with a rhodamine-goat anti-mouse IgG_{2a} antibody at 37°C are heavily labelled, but are devoid of antibody labelling at 4°C .

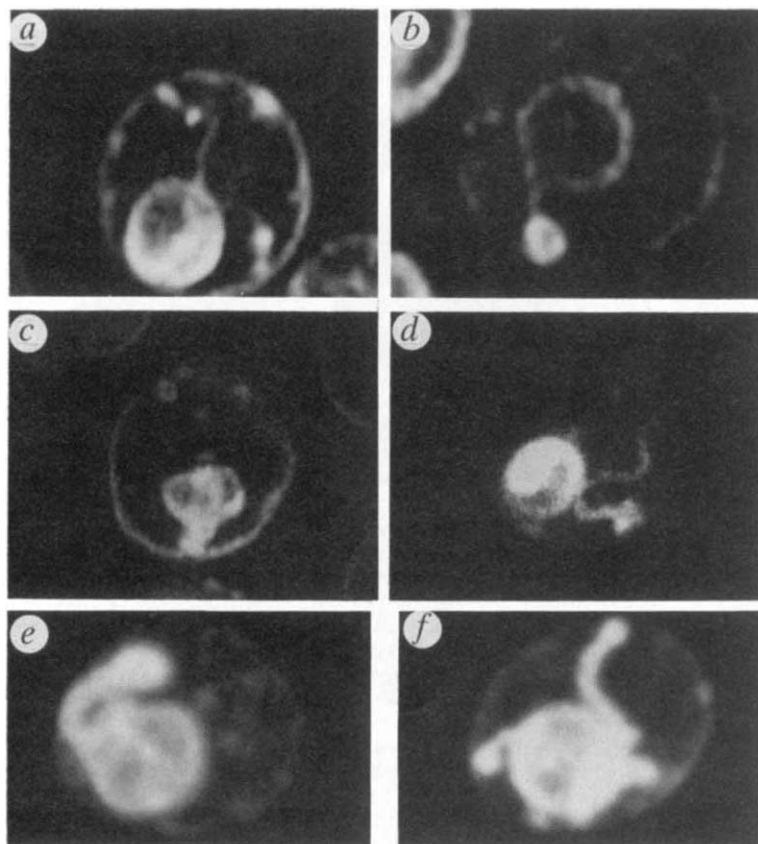
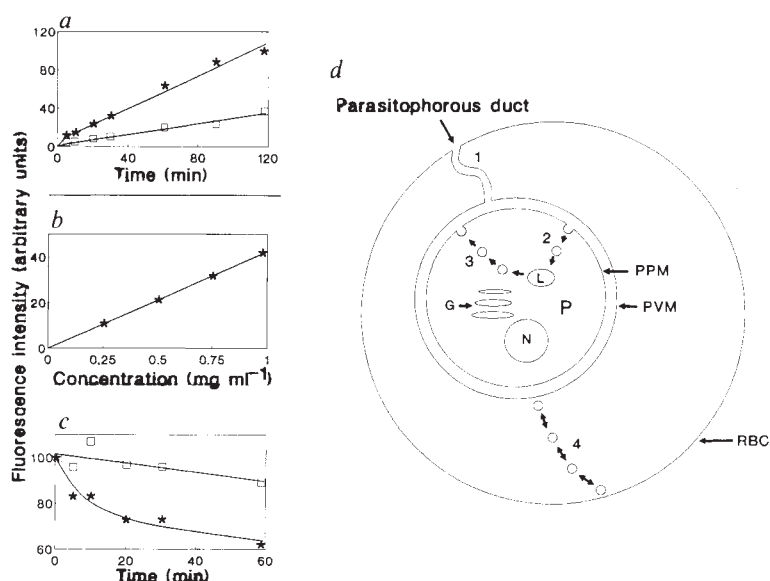


FIG. 3 Confocal fluorescence images of living, trophozoite-stage infected red blood cells incubated with: *a-d*, NBD-labelled phospholipids for 30 min at 37°C , or *e-f*, highly charged, carboxylate-modified, red fluorescent 30-nm latex FluoSpheres (Molecular Probes) for 1 h at 37°C . The spatial and vertical (*z*-axis) resolution of the confocal microscope under the conditions used in these experiments was ≈ 0.5 microns. *a*, C_6 -NBD-PC: the parasite is heavily labelled compared with the plasma membrane of the erythrocyte owing to the density and proximity of labelled intraparasitic membranes and organelles. Three tubular membranes, which we refer to as parasitophorous ducts, interconnect the parasite vacuolar and erythrocyte membranes. *b*, C_6 -NBD-PC: one duct connects the parasite vacuolar and erythrocyte membranes. It is not possible to say whether the vesicular structure at the end of the membranous tubule is one opening of the parasitophorous duct. *c*, C_6 -NBD-PE: one duct interconnects the parasite vacuolar and erythrocyte membrane. *d*, C_6 -NBD-ceramide: two tubular membrane projections extend from the parasitophorous vacuolar membrane (PVM), one of which appears to be continuous with the erythrocyte plasma membrane. *e* and *f*, Cells were incubated for 1 h at 37°C with carboxylate-modified fluorescent latex beads diluted 200 times with PBS which contained 1% bovine serum albumin and 3% of 40K dextran. Ducts projecting from the PVM, which traverse the erythrocyte cytoplasm, are clearly visible. On the basis of the biochemical uptake measurements, we suggest that these membranous tubules provide the path through which serum macromolecules diffuse into the aqueous compartment surrounding the parasite. The parasitophorous duct is likely to be distinct from Mauer's clefts, which are believed to be extensions of the PVM involved in the outward transport of proteins from the malaria parasite to the erythrocyte membrane¹⁷.

FIG. 4 *a*, Typical time course for accumulation of rhodamine-labelled dextran (10K) in the parasite of trophozoite-stage infections. Cells were incubated in tissue culture with rhodamine-dextran (1 mg ml^{-1}) at 37°C (*). For each time point, cells were taken from culture, washed 4 times at 4°C to remove external rhodamine-dextran and the fluorescence intensity in the parasite compartment quantified by confocal fluorescence image microscopy. Each fluorescence value is the average from 70–100 infected cells. Fluorescence values were obtained at fixed photomultiplier tube gains and are expressed in arbitrary units. To deplete ATP ($\square$), infected cells were incubated in glucose-free tissue culture medium, supplemented with 50 mM deoxyglucose and 5 mM NaN_3 , for 30 min at 37°C . *b*, Accumulation of rhodamine-dextran in the parasite of trophozoite-stage infections as a function of dextran concentration. Cells were incubated with increasing concentrations of rhodamine-dextran for 1 h at 37°C , washed 4 times at 4°C and the fluorescence in each compartment quantified by CFIM. Each value represents the average fluorescence from 70–100 infected cells. *c*, Infected cells were incubated for 30 min with rhodamine-dextran (1 mg ml^{-1}) at 37°C . A portion of the cells was washed and examined by CFIM at 4°C to obtain the 100% value (time, zero). The rest of the cells were washed at 37°C , and the amount of dextran remaining associated with the parasite quantified by CFIM as a function of time (asterisks). Each point is the average value measured for the parasite in 100 infected cells. If ATP is depleted during the 30 min incubation with rhodamine-dextran, exocytosis is inhibited and the level of parasite fluorescence remains relatively constant ($\square$). *d*, A schematic model of communication pathways in infected cells. (1) Serum macromolecules passively diffuse into the parasitophorous duct; (2) endocytosis of macromolecules by the parasite (P); (3) exocytosis from



parasite lysosomes (?), (L); (4) vesicles in the erythrocyte cytosol could be derived from endocytosis of the erythrocyte membrane, parasite exocytosis to the erythrocyte or a combination of both. The mechanism by which endocytic and exocytic vesicles cross the parasitophorous vacuolar membrane (PVM) is unknown. PPM, parasite plasma membrane; N, nucleus; G, Golgi.

The kinetics of uptake of rhodamine-dextran (10K) into the parasite at 37°C are shown in Fig. 4*a*. An uptake rate, characteristic of endocytosis, was obtained⁶. Accumulation of rhodamine-dextran was inhibited by depletion of ATP. The accumulation of rhodamine-dextran in the parasite of trophozoite-stage cells was also measured as a function of rhodamine-dextran concentration (Fig. 4*b*), where a linear relationship, typical of fluid-phase endocytosis⁶, was observed.

No direct evidence is available regarding the ability of the intracellular parasite to exocytose. Studies describing the appearance, post-infection, of parasite lipids^{2,7}, proteins^{8–10} and transferrin receptors¹¹ in the erythrocyte membrane suggest that exocytic pathways exist. Figure 4*c* presents evidence for parasite exocytosis. Parasites in infected cells were loaded with rhodamine-dextran for 30 min at 37°C and washed to remove free rhodamine-dextran. A rapid decrease in fluorescence was observed during the first 10 min after removal of unincorporated rhodamine-dextran. Exocytosis was inhibited when ATP was depleted during the pulse with rhodamine-dextran at 37°C , with the level of parasite fluorescence remaining constant.

Our results are evidence for passive diffusion of macromolecules in serum through a newly identified parasitophorous duct, with subsequent endocytosis of these macromolecules by the parasite (Fig. 4*d*). This system would

provide the parasite with an energy-efficient pathway to obtain nutrients and protein-bound fatty acids and lipids from the serum. The properties of this parasitophorous duct would be consistent with the results of Elford *et al.*¹², who suggested that young trophozoites have an uninterrupted access to the extra-erythrocytic fluid phase. Our data show that the view of the parasite as surrounded by a continuous, sealed vacuole¹ needs to be modified. At present, it is not clear whether the parasitophorous duct represents an invagination of the erythrocyte membrane which persists from the time of merozoite invasion or a membrane-bound pathway formed at a later time in the parasite's life cycle.

Our observation that an IgG antibody has access to the parasite may explain some hitherto enigmatic immunological observations. IgG antibodies¹³ and crisis-form factor¹⁴ from human immune serum retard the *in vitro* growth of *P. falciparum*. The parasitophorous duct may provide the pathway for these molecules to reach the 'intracellular' parasite. Until now, surface antigens on trophozoites and schizonts have not been considered as important for development of blood-stage malaria vaccines. IgG antibodies elicited by vaccination, specific for antigens on the parasite plasma membrane, may be parasitocidal when coupled to toxins or other agents that compromise the biochemical viability of the malaria parasite. □

Received 28 March; accepted 25 July 1991.

1. Aikawa, M., Miller, L. H., Johnson, J. & Rabbage, J. *J. Cell Biol.* **77**, 72–81 (1978).
2. Hsiao, L., Howard, R. J., Aikawa, M. & Taraschi, T. F. *Biochem. J.* **274**, 121–132 (1991).
3. Hunter, J. A., Hurlley, S. M., Murray, M. & Taraschi, T. F. *Biophys. J.* **59**, 442a (1991).
4. Kuusimäen, E. & Saraste, J. *Meth. Cell Biol.* **32B**, 257–273 (1989).
5. Grellier, P., Rigomier, D., Clavey, V., Fruchart, J.-C. & Schrevel, J. *J. Cell Biol.* **112**, 267–277 (1991).
6. Swanson, J. A., Yirinec, B. D. & Silverstein, S. C. *J. Cell Biol.* **100**, 851–859 (1985).
7. van der Schaft, P. H. *et al. Biochim. Biophys. Acta* **901**, 1–14 (1986).
8. Howard, R. J. *et al. J. Cell Biol.* **103**, 1269–1277 (1986).
9. Howard, R. J. *et al. J. Cell Biol.* **104**, 1269–1280 (1987).
10. Taylor, D. W. *et al. Molec. Biochem. Parasitol.* **25**, 165–174 (1987).

11. Rodriguez, M. & Jungery, M. *Nature* **324**, 388–391 (1986).
12. Elford, B. C., Haynes, J. D., Chulay, J. D. & Wilson, R. J. M. *Molec. Biochem. Parasitol.* **16**, 43–60 (1985).
13. Jepsen, S. *Scand. J. Immun.* **18**, 567–571 (1983).
14. Jensen, J. B., Boland, M. T. & Akood, M. A. *Science* **216**, 1230–1233 (1982).
15. Green, T. J. *et al. Am. J. Trop. Med. Hyg.* **34**, 24–30 (1985).
16. Trager, W. & Jensen, J. B. *Science* **193**, 673–675 (1976).
17. Barnwell, J. W. *Blood Cells* **16**, 379–395 (1990).

ACKNOWLEDGEMENTS. The expert technical assistance of W. L. Chung is acknowledged. Supported by the US Public Health Service (T.F.T.) and the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases (T.F.T.).

Targeting of transmembrane and GPI-anchored forms of N-CAM to opposite domains of a polarized epithelial cell

Sharon K. Powell, Bruce A. Cunningham*,
Gerald M. Edelman* & Enrique Rodriguez-Boulan

Department of Cell Biology and Anatomy,
Cornell University Medical College, 1300 York Avenue, New York,
New York 10021, USA

* Department of Developmental and Molecular Biology,
Rockefeller University, New York, New York 10021, USA

THE calcium-independent neural cell adhesion molecule N-CAM is expressed transiently during development in many tissues, including epithelia¹. The three naturally occurring principal isoforms^{2,3} of N-CAM differ in the way in which they associate with the membrane and in their cytoplasmic domains⁴. These isoforms are generated by developmentally regulated alternative splicing of a single gene⁴⁻⁶: the large cytoplasmic domain (ld) form (relative molecular mass 180,000 (M_r 180K)) is specific for post-mitotic neurons; the 120K small cytoplasmic domain (ssd) and 140K small surface domain (sd) forms also occur on other cell types⁷. One function of the different isoforms could be to specify cellular localization; for example, glycosyl phosphatidyl inositol (GPI)-membrane anchoring acts as a targeting signal for expression on the apical surface of polarized epithelial cells^{8,9}. Neurons and epithelial cells may use similar mechanisms for polarizing their plasma membrane proteins^{10,11}. We have therefore investigated the targeting of GPI-anchored (ssd N-CAM, 120K) and transmembrane forms of N-CAM (sd N-CAM, 140K; ld N-CAM, 180K) by comparing the expression of each after transfection of the appropriate complementary DNAs into polarized epithelial cells. We find that isoforms with alternative modes of membrane association are targeted to different surfaces of polarized epithelial cells: ssd N-CAM is expressed on the apical surface, whereas sd and ld N-CAM are expressed on the basolateral surface. These results suggest that the different isoforms of N-CAM determine their own diverse cellular destinations. They also support the hypothesis that the GPI anchor acts as an apical targeting signal in epithelia.

After transfection of the cDNAs for chicken ssd, sd and ld N-CAM into the canine kidney epithelial cell line MDCK we determined the localization of the isoforms by indirect immunofluorescence. The ssd N-CAM is expressed apically (Fig. 1a, b, g, h), and both sd (Fig. 1c, d) and ld (Fig. 1e, f) N-CAM are expressed basolaterally, indicating that the mode of membrane association correlates with localization on the apical or basolateral surface. To quantitate the relative expression of each form on the apical and basolateral surfaces, clonal lines transfected with the N-CAM cDNAs were grown on Transwell filters, biotinylated from either the apical or the basolateral surface¹², and immunoprecipitated with anti-N-CAM antibody. More than 95% of ssd N-CAM was labelled apically, and more than 95% of sd and ld N-CAM were labelled from the basolateral surface (Fig. 2).

Intracellular sorting and polarized delivery have been demonstrated for several proteins in MDCK cells¹³. We investigated cell-surface delivery of clones expressing ssd or sd N-CAM by a pulse-chase/biotin targeting assay which detects only N-CAM that has reached the cell surface at the time of biotinylation. The ssd N-CAM appeared directly on the apical surface (Fig. 3, upper panel). Targeting of sd (Fig. 3, lower panel) and ld N-CAM (data not shown) to the basolateral surface was also direct. These experiments indicate that the isoforms directed to opposite surfaces are segregated intracellularly, presumably into

different post-trans-Golgi network vesicles that are targeted vectorially to the appropriate membrane domain. As the ectodomain (682 amino acids) is common to all three forms, our results strongly support the existence of targeting signals in anchor or cytoplasmic domains of the protein.

The predominant cell adhesion molecule expressed in adult epithelial tissues such as liver and kidney is the calcium-dependent molecule L-CAM (E-cadherin, uvomorulin)^{1,14}. Previous

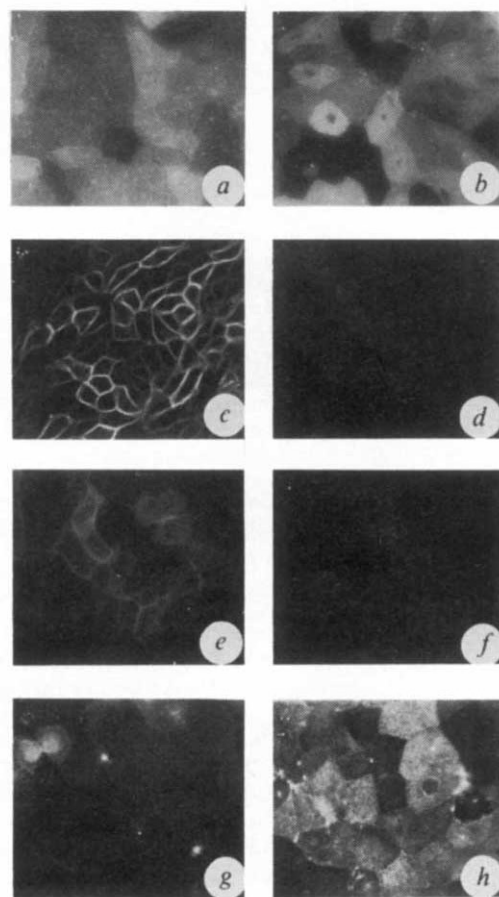


FIG. 1 Indirect immunofluorescence of cells transfected with ssd, sd, and ld N-CAM. The cDNAs for chicken sd and ssd N-CAM were subcloned on *Bam*HI fragments and ligated into the *Bgl*II site of an expression vector using a Rous sarcoma virus promoter²⁴. The gene for chicken ld N-CAM was under the control of the thymidine kinase promoter²⁵. These plasmids were co-transfected with pMV6neo into MDCK cells by the calcium phosphate method²⁶. After selection in 500 μ M ml⁻¹ G418, a sample from the total G418-resistant population transfected with each plasmid was grown on coverslips and fixed with 2% paraformaldehyde. N-CAM was detected by incubation with a 1:300 dilution of rabbit anti-chicken N-CAM followed by 1:100 rhodamine goat anti-rabbit IgG. In a, c and e, cells were permeabilized with 0.1% saponin before antibody incubation. Permeabilization of the cells allows access of the antibody to the cytoplasm and to the basal and lateral surfaces. In b, d and f–h cells were exposed to antibody immediately after fixation, without permeabilization. In the absence of permeabilization, only the apical surface is accessible to antibody. a, b, g and h, Cells expressing ssd N-CAM; c and d, sd N-CAM; e and f, ld N-CAM. For ssd-expressing cells, the pattern is very similar in permeabilized (a) and non-permeabilized (b) samples. In contrast, little staining for sd (d) and ld (f) could be detected on the apical surface (non-permeabilized cells); in permeabilized cells, staining was concentrated along the lateral surfaces for both transmembrane forms (c, e). In g, cells expressing ssd N-CAM were treated for 1 h with 5 units ml⁻¹ phosphoinositol-specific phospholipase C (PI-PLC; gift from M. Low, Columbia University) in DMEM before staining; cells in h were incubated in the same way but no PI-PLC was added. Staining for another apical marker that is not GPI-anchored (integral membrane glycoprotein gp114; ref. 13) or for transmembrane forms of N-CAM is not affected by treatment with PI-PLC (data not shown). All samples plus and minus saponin and plus and minus PI-PLC were photographed and developed at identical exposures.

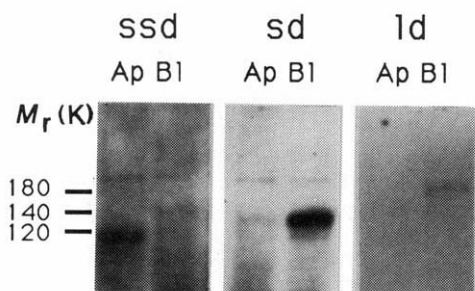


FIG. 2 Domain selective biotinylation of cells expressing ssd, sd, and ld N-CAM. Cells grown on Transwell filters were incubated for 30 min at 4 °C with 0.5 mg ml⁻¹ sulfo-NHS-biotin (Pierce) from either the apical (Ap) or basolateral (Bl) chamber as described^{2,27}, then extracted in lysis buffer A (150 mM NaCl, 20 mM Tris-HCl, pH 8.0; 5 mM EDTA, 1% Triton X-100, 0.2% BSA) and immunoprecipitated with 2 µl of rabbit anti-N-CAM antibody (which recognizes all three N-CAM isoforms). Immune complexes were precipitated with protein A-Sepharose (Pharmacia) and washed once with buffer A, three times with buffer A + 0.1% SDS, three times in buffer C (500 mM NaCl, 20 mM Tris, pH 8.0, 0.5% Triton X-100, 0.2% BSA), and once with 50 mM Tris, pH 8.0. After electrophoresis on 10% polyacrylamide gels and transfer to nitrocellulose, biotinylated proteins were detected by incubation with ¹²⁵I-streptavidin.

experiments in which N-CAM or L-CAM were introduced into non-epithelial cells by transfection demonstrated that expression of these cell-adhesion molecules affected the morphology of the monolayer, resulting in the formation of an 'epithelioid' sheet^{15,16}. N-CAM-expressing MDCK cells do not exhibit altered morphology. Untransfected MDCK cells and cells expressing N-CAM incubated in low calcium medium to disrupt E-cadherin-mediated adhesion display no differences in transepithelial resistance of the monolayers (data not shown), indicating that N-CAM cannot substitute for E-cadherin under these conditions. E-cadherin associates with a basolateral membrane-cytoskeletal complex which may be important in driving epithelial polarization^{17,18}; perhaps N-CAM lacks the ability to interact with such complexes.

Our results show that the GPI and transmembrane forms of N-CAM are efficiently targeted to opposite surfaces of the polarized epithelial line MDCK. They strongly support the

hypothesis that GPI anchors serve as targeting signals for apical localization, and in addition indicate that the different modes of membrane anchoring for a single protein may direct protein localization.

It has been suggested that homophilic interactions of cell adhesion molecules on adjoining cells is responsible for their localization to the lateral surfaces of epithelial cells¹⁹. The ssd, sd and ld N-CAM are all identical from their amino termini up to amino 682, just adjacent to the membrane-attachment domain⁴, and have identical adhesion properties²⁰. Our results therefore suggest that ectodomain interactions are not sufficient for localization to the lateral surface and that the GPI-targeting signal for the apical surface may be dominant over such interactions.

It has been suggested that neurons maintain segregated membrane domains analogous to the apical and basolateral domains of polarized epithelial cells¹⁰. Although N-CAM generally seems to be expressed over the entire plasma membrane of neuronal cells, it is interesting to note that the ld form can be concentrated (but not exclusively expressed) in growth cones, at sites of cell-cell contact, and at postsynaptic densities²¹⁻²³. Because antibodies against the ssd and sd forms also react with the ld form (owing to their common ectodomain), the differential localization of N-CAM isoforms in a single neuron has not been clearly addressed. Given the distinct localization of N-CAM isoforms described here, it will be of interest further to compare targeting signals and membrane domains in neurons and epithelial cells. □

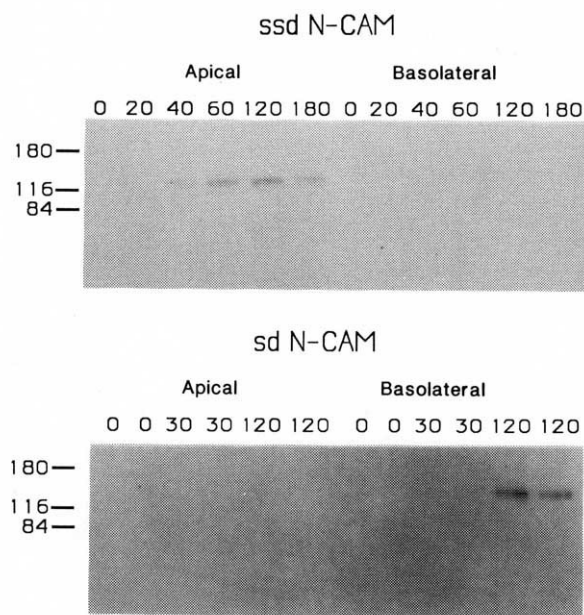


FIG. 3 Vectorial delivery of ssd and sd N-CAM to the cell surface. A biotin targeting assay was used to follow the delivery of ssd and sd N-CAM to the membrane²⁷. Cells were pulsed with Expre³⁵S protein labelling mix (NEN) for 30 min. At varying times of chase in DMEM containing 50 mM HEPES and 0.2% BSA, samples were placed on ice and surface-labelled with sulfo-NHS-biotin (Pierce) from either the apical or basolateral chamber as described for Fig. 2. Cells were lysed in buffer A and precipitated with rabbit anti-chicken N-CAM and protein A-Sepharose beads as described in the legend to Fig. 2. The antigen was eluted from the immune complex pellet by boiling the beads in 10 µl 10% SDS for 5 min, diluted with 500 µl buffer A, and precipitated with streptavidin agarose beads (Pierce) by incubation overnight at 4 °C. By this technique, only the labelled antigen that has reached the cell surface at the time of biotinylation is precipitated. Samples were then electrophoresed on 10% polyacrylamide gels and processed for autoradiography. Time of chase in minutes at biotinylation is indicated above each lane, and the numbers to the left are M_r ($\times 10^{-3}$) markers.

Received 13 May; accepted 18 July 1991.

- Edelman, G. M. *A. Rev. Cell Biol.* **2**, 81-116 (1986).
- Klein, G., Langeegger, M., Goridis, C. & Ekblom, P. *Development* **102**, 749-761 (1988).
- Thiery, J.-P., Duband, J.-L., Rutishauser, U. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 6737-6741 (1982).
- Cunningham, B. A. *et al. Science* **236**, 799-806 (1987).
- Hemperly, J. J., Edelman, G. M. & Cunningham, B. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9822-9826 (1986).
- Owens, G. C., Edelman, G. M. & Cunningham, B. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 294-298 (1987).
- Cunningham, B. A. & Edelman, G. M. in *Morphoregulatory Molecules* (eds Edelman, G. M., Cunningham, B. A. & Thiery, J.-P.) (Wiley, New York, 1990).
- Lisanti, M., Caras, I. P., Davitz, M. A. & Rodriguez-Boulan, E. *J. Cell Biol.* **109**, 2145-2156 (1989).
- Lisanti, M. P. & Rodriguez-Boulan, E. *Trends biochem. Sci.* **15**, 113-118 (1990).
- Dotti, C. G. & Simons, K. *Cell* **62**, 63-72 (1990).
- Dotti, C. G., Parton, R. G. & Simons, K. *Nature* **349**, 158-161 (1991).
- Sargiacomo, M., Lisanti, M., Graeve, L., Le Bivic, A. & Rodriguez-Boulan, E. *J. Memb. Biol.* **107**, 277-286 (1989).
- Le Bivic, A., Sambuy, Y., Mostov, K. & Rodriguez-Boulan, E. *J. Cell Biol.* **110**, 1533-1539 (1990).
- Takeichi, M. *A. Rev. Biochem.* **59**, 237-252 (1990).
- Matsuzaki, F. *et al. J. Cell Biol.* **110**, 1239-1252 (1990).
- Mege, R.-M. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 7274-7278 (1988).
- Nelson, W. J., Shore, E. M., Wang, A. Z. & Hammertson, R. W. *J. Cell Biol.* **110**, 349-357 (1990).
- McNeill, H., Ozawa, O., Kemler, R. & Nelson, W. J. *Cell* **62**, 309-316 (1990).
- Fleming, T. P. & Johnson, M. H. *A. Rev. Cell Biol.* **4**, 459-485 (1988).
- Pizzey, J. A., Rowlett, L. H., Barton, C. H., Dickson, G. & Walsh, F. S. *J. Cell Biol.* **109**, 3465-3476 (1989).
- Pollerberg, G. E., Burridge, K., Krebs, K. E., Goodman, S. R. & Schachner, M. *Cell Tissue Res.* **250**, 227-236 (1987).
- Letourneau, P. C. & Shattuck, T. A. *Development* **105**, 505-519 (1989).
- Persohn, E., Pollerberg, G. E. & Schachner, M. *J. comp. Neurol.* **288**, 92-100 (1989).
- Moore, H.-P. & Kelly, R. B. *J. Cell Biol.* **101**, 1773-1781 (1985).
- Murray, B. A., Hemperly, J. J., Prediger, E. A., Edelman, G. M. & Cunningham, B. A. *J. Cell Biol.* **102**, 189-193 (1986).
- Graham, F. & Van der Eb, A. *Virology* **52**, 456-467 (1973).
- Le Bivic, A., Real, F. X. & Rodriguez-Boulan, E. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9313-9317 (1989).

ACKNOWLEDGEMENTS. We thank M. Lisanti for his help in the initiation of this project. This work was supported by grants from ACS and the NIH.

Failure of wild-type or a mutant form of protein kinase C- α to transform fibroblasts

Christoph Borner*, Ireos Filipuzzi†, I. Bernard Weinstein & Roland Imber†

Comprehensive Cancer Center and Institute for Cancer Research, Columbia University, 701 West 168th Street, New York, New York 10032, USA

† Laboratory of Molecular Tumorbiology, University Clinic Medical School, 4031 Basel, Switzerland

* To whom correspondence should be sent at Institute of Biochemistry, University of Lausanne, Ch. des Boveresses 155, CH-1066 Epalinges, Switzerland

A MUTANT form of the α -isoform of protein kinase C (PKC) was recently isolated from an ultraviolet radiation-induced murine fibrosarcoma cell line and reported to transform mouse BALB/c 3T3 fibroblasts on transfection¹. Four point mutations in the regulatory domain were assumed to be responsible for its oncogenicity and unusual preference for membrane localization. Here, we report that overexpression of the reported mutant PKC α complementary DNA in three fibroblast cell lines, including BALB/c 3T3, does not enable these cells to grow in soft agar or nude mice. In addition, this mutant PKC α form seems to be indistinguishable from the wild-type PKC α with respect to its dependence on cofactors, phorbol ester binding, subcellular distribution and its effects on growth and morphology. These results fail to confirm the previous study¹ and indicate that overexpression of either the

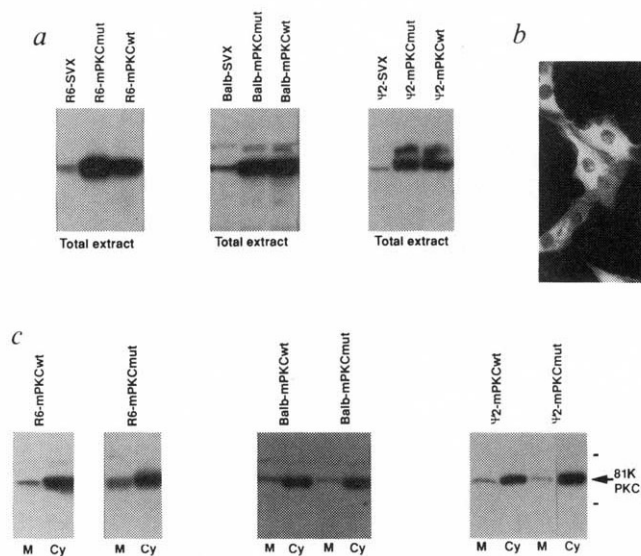
wild-type or the reported mutant form of PKC α does not transform rodent fibroblasts.

The reported mutant PKC α (UV25; hereafter called mPKCmut) and its wild-type counterpart (mPKCwt)¹ were individually overexpressed in rat 6 (R6), mouse BALB/c 3T3 and mouse Ψ 2 fibroblasts. These cell lines were chosen for the following reasons: (1) R6 cells can be partially or completely transformed on overexpression of wild-type PKC β 1 or various oncogenes, respectively^{2,3}; (2) BALB/c 3T3 are the cells originally used to show the transforming potential of mPKCmut (ref. 1); and (3) Ψ 2 cells are extremely sensitive to transformation by oncogenes⁴.

Cell lines transfected with mPKCmut or mPKCwt expressed 20–40-fold higher amounts of the corresponding mutant or wild-type PKC α protein than the levels of PKC α present in the respective vector control cells (Fig. 1a, Table 1). In all cases, the increases in mutant and wild-type PKC α protein correlated with increases in phorbol ester binding and protein kinase activities (Table 1). Detailed analysis on PKC activities revealed that both the mutant and the wild-type PKC α activities were equally dependent on phosphatidylserine (PS) and Ca²⁺ and stimulated by diacylglycerol (DAG) (Table 1). Most important, the unstimulated activity of the mutant PKC α was comparable to that of wild-type PKC α (1–2% of the DAG-stimulated activity) indicating that the former does not exhibit a cofactor-independent, constitutive activity. When analysing the subcellular distribution of the two PKC forms, we reproducibly found that the mutant protein (Fig. 1c, Table 1) displayed a similar extent of membrane association as wild-type PKC α (15–30%). In addition, immunofluorescence on R6-mPKCmut cells illustrated that the mutant PKC α protein was mainly located in the cytoplasm (Fig. 1b). This result is in disagreement

FIG. 1 Stable expression of mutant and wild-type murine PKC α in fibroblasts. Protein immunoblots of the R6, BALB/c and Ψ 2 cell lines showing the expression of mutant and wild-type PKC α (81K) in total extracts (a) and cytosolic (Cy)/membrane (M) fractions (c). Immunofluorescence of mutant PKC α in whole R6-mPKCmut cells (b). Dashes mark the positions of proteins with apparent M_r of 93K and 68K, respectively.

METHODS. mPKCmut (UV25) and mPKCwt cDNAs were subcloned into the retroviral vector pZIP-Neo SV(X)¹ (ref. 5). Early passage R6 embryo fibroblasts, mouse BALB/c 3T3 (ATCC, CCL 163) or mouse Ψ 2/NIH 3T3 fibroblasts were transfected with the SVX-mPKCmut, SVX-mPKCwt or SVXneo plasmids, using the calcium phosphate precipitation technique⁶. Following selection in media containing 200 μ g ml⁻¹ G418, 20 clones of each transfectant were picked, expanded and analysed by northern blot analysis for the expression of the respective construct. G418-resistant R6, BALB/c and Ψ 2 cells expressing large amounts of the correctly sized mutant PKC α mRNA (R6-mPKCmut, BALB-mPKCmut, Ψ 2-mPKCmut) or large amounts of wild-type PKC α mRNA (R6-mPKCwt, BALB-mPKCwt, Ψ 2-mPKCwt) and SVX control cells (R6-SVX, BALB-SVX, Ψ 2-SVX) were chosen for further studies. Cells were grown in Dulbecco's modified Eagle medium (DMEM, Gibco) supplemented with 10% calf serum (Flow). The medium was changed every 3 days; starving and growth to postconfluence were strictly avoided. The cells were collected when subconfluent. For protein and enzyme assays, the cells were washed three times in PBS, and either immediately lysed in preheated (95 °C) extraction buffer (20 mM Tris-HCl, pH 7.5, 2 mM EDTA, 2 mM EGTA, 6 mM β -mercaptoethanol, 2 μ g ml⁻¹ aprotinin, 10 μ g ml⁻¹ leupeptin) containing 1% SDS (total extracts) or scraped into ice-cold extraction buffer, lysed by sonication and centrifuged at 100,000g for 1 h, yielding the soluble cytosol and the membrane pellet. The membrane pellet was solubilized in preheated extraction buffer containing 1% SDS. The protein content in these extracts was determined according to the method of Lowry. Immunoblot analyses of the cytosolic, membrane and total extracts were performed exactly as described⁷. Mutant and wild-type PKC α were immunodetected using primary antibodies specific for PKC α (Amersham, MC5)⁷ and ¹²⁵I-labelled secondary antibodies (sheep anti-mouse, Amersham). For immunocytochemical analysis, subconfluent R6 cells were fixed, permeabilized, stained with the PKC α specific monoclonal antibody M6 and finally with fluorescein-conjugated second antibody, as previously described⁸. Total RNA was isolated from the various cell lines exactly as described⁷. Hybridization to RNA blots was done with a ³²P-labelled murine PKC α probe (5' PstI-BglII fragment) under high stringency conditions¹.



with the reported data which claimed that 80% of the mutant form is associated with the membrane¹.

Next, we compared the morphology and growth properties of the different cell lines. The results with R6 cells (Fig. 2) were the same as with BALB/c and Ψ 2 cells. After treatment with 12-*O*-tetradecanoyl phorbol 13-acetate (TPA), control cells showed a slight change in morphology with a return to normal appearance by 24 h. By contrast, as previously shown for PKC β I-overproducing cells (R6-PKC3)², R6-mPKCmut and R6-mPKCwt cells displayed a persistent and exaggerated morphological change in response to TPA. This finding clearly shows that both the mutant and wild-type PKC α , are functional *in vivo* and their activation by TPA leads to a protracted cellular response. We did not, however, notice any difference in the altered morphology between mPKCmut and mPKCwt cells. In addition, none of the mPKCmut cells exhibited a TPA-like morphology in the absence of TPA which indicates that the mutant PKC α enzyme is not constitutively activated.

We also compared the growth rates of the R6 cell derivatives because it was previously shown that R6 cells overexpressing PKC β I exhibited increased growth and a higher saturation density². R6-mPKCmut and R6-mPKCwt cells were indistinguishable with respect to growth rate and saturation density (Fig. 2). They both grew slower than control cells, especially in the presence of TPA, and their saturation densities were lower in the absence of TPA, but reached control levels in the presence of TPA.

To assess the possibility that the fibroblasts overexpressing mutant PKC α may nevertheless exhibit anchorage-independent growth properties, we tested the ability of the transfectants to form colonies in soft agar. As previously shown, R6 cells overexpressing PKC β I formed small colonies in soft agar and the colony sizes and efficiencies were enhanced when TPA was included in the medium² (Fig. 2). By contrast, R6 as well as BALB/c 3T3 cells overexpressing large amounts of wild-type or mutant PKC α failed to grow in soft agar and persisted as

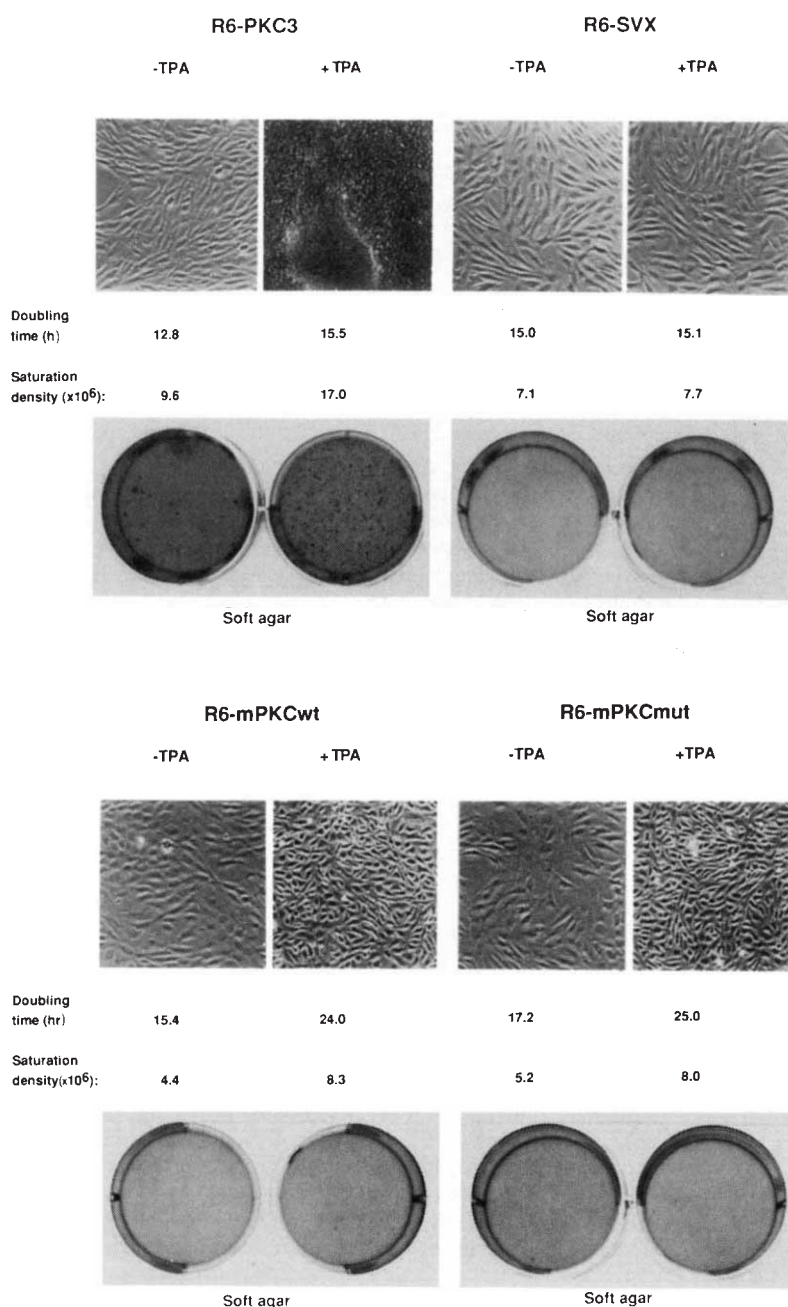


FIG. 2 Morphological responses and growth properties. Photographs of the cells were taken at magnification $\times 100$ (for TPA-treated R6-PKC3 cells, $\times 40$) after 6 days of growth in the presence or absence of 100 ng ml^{-1} TPA. Photographs of the soft agar colonies were taken after 30 days of growth; magnification $\times 40$. The doubling times in the absence or presence of 100 ng ml^{-1} TPA relate to the initial exponential phase of cell growth, the saturation densities to the number of cells per 35-mm plate on day 28. The saturation density of R6-PKC3 cells was determined at day 12 as these cells peel off the plate thereafter, especially in the presence of TPA. METHODS. Cells were seeded at a density of 10^4 cells per 35-mm plate (diameter) in 3 ml DMEM plus 10% calf serum (without G418). Cells were grown in the absence or presence of TPA (100 ng ml^{-1}) and counted on triplicate plates as previously described² with the modification that the medium was changed every 2 days. To assess growth in soft agar, 2×10^4 cells were suspended in the absence or presence of 100 ng ml^{-1} TPA in 2 ml 0.3% Bacto-agar (Difco Laboratories) in DMEM containing 10% fetal calf serum (without G418) and overlaid above a layer of 5 ml 0.5% agar in the same medium on 35-mm Petri dishes. The cells were then overlaid with 2 ml DMEM plus 10% FCS every 4 days. At the end of 30 days, colonies were stained and counted as described².

TABLE 1 [³H]Phorbol 12,13-dibutyrate binding and PKC activity in mPKCwt- and mPKCmut-transfected cells

	Specific PKC activity* pmol min ⁻¹ (mg protein) ⁻¹ (%)					[³ H]PDBu† binding pmol mg ⁻¹		PKCα‡ protein	%M§
	Complete	-DAG	-Ca ²⁺	-PS	+EGTA	Fold	Fold	Fold	
R6-SVX	181 (100)	103 (57)	51 (28)	<10	<10	1	1.3	1	15
R6-mPKCwt	4842 (100)	2576 (53)	1598 (33)	97 (2)	44 (1)	27	15.7	12	18
R6-mPKCmut	6804 (100)	3395 (50)	2048 (30)	151 (2)	75 (1)	38	19.9	15	30
Balb-SVX	832 (100)	508 (61)	285 (34)	25 (3)	12 (1)	1	3.2	1	17
Balb-mPKCwt	14774 (100)	8584 (58)	4583 (31)	443 (3)	133 (1)	18	26.3	8	18
Balb-mPLCmut	10503 (100)	5912 (56)	3707 (35)	266 (3)	158 (2)	13	21.3	7	25
Ψ2-SVX	310 (100)	185 (60)	102 (33)	<10	<10	1	2.1	1	17
Ψ2-mPKCwt	8681 (100)	4502 (52)	3120 (36)	263 (3)	79 (1)	28	25.2	12	19
Ψ2-mPKCmut	8063 (100)	4345 (54)	2811 (35)	235 (3)	92 (1)	26	23.2	11	24

Cytosolic PKC was partially purified by DEAE-Sephacel chromatography as described⁹. Fractions from a NaCl-gradient elution (0–0.3 M) exhibiting PKC activity were pooled, concentrated by dialysis against 30% sucrose and reassayed for PKC. PKC activity was measured using an EGF-receptor (EGF-R) peptide as substrate, as described^{2,10,11}. The complete reaction mixture (250 μl) contained 20 mM Tris-HCl, pH 7.5, 10 mM Mg(NO₃)₂, 100 μg phosphatidylserine (PS), 10 μg 1,2-diolein (DAG), 24 μM [γ-³²P]ATP (0.5 μCi), 400 μM CaCl₂, 100 μM EGTA, 100 μM EGF-R peptide. For the assay of unstimulated activity (+EGTA), 0.5 mM EGTA was included in the reaction mixture instead of PS/DAG and Ca²⁺. Following incubation for 20 min at 32 °C, a 100 μl aliquot was removed from the reaction mixture and applied to P-81 paper (Whatman), washed and counted. Phorbol ester binding was done on cytosolic fractions as previously described¹². The values reflect total binding minus nonspecific binding measured in the presence of 2 μM TPA.

* Specific PKC activity of the pooled fractions eluted from a DEAE-Sephacel column loaded with the cytosol of the respective cell line. The complete reaction system contained 1,2-diolein (DAG), CaCl₂ (Ca²⁺) and phosphatidylserine (PS). In some assays, the cofactors were sequentially removed in the order of DAG, Ca²⁺, PS. Basal activity was measured in the presence of EGTA (+EGTA) lacking DAG, Ca²⁺ and PS. Values in brackets indicate the percentage of the maximally stimulated PKC activity (100%) obtained in the complete system.

† Specific [³H]phorbol 12,13-dibutyrate-binding to cytosolic proteins.

‡, § Mutant and wild-type PKCα protein in total extracts (‡) or in cytosol/membrane fractions (§) were quantitated from autoradiographs of immunoblots using a laser densitometer (Molecular Dynamics). Values were taken from the linear range of a ¹²⁵I-standard curve. %M reflects the ratio of membrane-bound/total in percentage. Similar results were obtained from three independent experiments.

single cells comparable to vector control cells, even in the presence of TPA (Fig. 2). In addition, inoculation of R6-mPKCmut as well as R6-mPKCwt cells into nude mice did not lead to tumour formation even after 2 months (data not shown) although R6 cells overexpressing PKCβ1 produced tumours at a slow rate².

These studies indicate that the reported mutated form of PKCα isolated from an ultraviolet radiation-induced fibrosarcoma cell line does not transform rat and mouse fibroblasts and functions exactly like a wild-type enzyme when stably overexpressed *in vivo* or assayed *in vitro*. It can therefore be safely argued that this point-mutated PKCα is not an oncogene. These data run counter to the finding of Megidish *et al.*¹ in all major points. We do not know the exact reason for this discrepancy but suspect that it may in part be related to the cell line used in the previous study. We learned in the course of our own experiments that untransfected normal BALB/c 3T3 cells easily acquire a transformed phenotype and grow in soft agar when extensively passaged in culture or held at postconfluence. Thus, the previous study¹ might have involved inadvertent selection of spontaneously transformed cells. □

Wee1⁺-like gene in human cells

Makoto Igarashi, Akihisa Nagata, Shigeki Jinno,
Kimihiko Suto & Hiroto Okayama*

Department of Molecular Genetics, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565, Japan

THE *wee1*⁺ gene is a mitotic inhibitor controlling the G2 to M transition of the fission yeast *Schizosaccharomyces pombe*¹ and encodes a protein kinase with both serine- and tyrosine-phosphorylating activities². We have cloned a human gene (*WEE1Hu*) similar to *wee1*⁺ by transcomplementation of a yeast mutant^{3,4}. *WEE1Hu* encodes a protein homologous to the *S. pombe wee1*⁺ and *miki*⁺ (a functionally redundant sibling of *wee1*⁺)⁵ kinases and effectively rescues a *wee1* mutation. We report here that overexpression of *WEE1Hu* in fission yeast generates very elongated cells as a result of inhibition of the G2–M transition in the cell cycle. In addition, we detected a 3-kilobase-long *WEE1Hu* messenger RNA in all the human cell lines we examined. We conclude that a *wee1*⁺-like gene exists and is expressed in human cells.

S. pombe defective in the *wee1*⁺ gene has a short G2 period and enters mitosis at a reduced cell size. When a mitotic inducer *cdc25*⁺ is overexpressed in a *wee1* mutant, cells die by prematurely induced mitosis⁶. To isolate a mammalian counterpart of the *wee1*⁺ gene, the *S. pombe* temperature-dependent lethal mutant *wee1-50 leu1-32 h*⁺ with an overproducing *cdc25*⁺ gene was transfected with a primary human fibroblast complementary DNA library⁷ together with the pAL19 yeast vector⁴. Among 3 × 10⁶ stably transfected colonies, 10 grew at the restrictive temperature and one cDNA clone (*WEE1Hu*) with strong complementation activity was recovered from one of the colonies.

The *WEE1Hu* cDNA was 2.8 kilobases (kb) long, excluding a poly(A) tail, and was judged as a full-length or nearly full-length clone from the size of the mRNA (about 3 kb; see below). Its nucleic-acid sequence and the amino-acid sequence of the encoded protein are shown in Fig. 1. The encoded protein is 432 amino-acids long with an estimated relative molecular mass

Received 15 April; accepted 28 June 1991.

- Megidish, T. & Mazurek, N. *Nature* **342**, 807–811 (1989).
- Housey, G. M. *et al. Cell* **52**, 343–354 (1988).
- Hsiao, W.-L. W., Housey, G. M., Johnson, M. D. & Weinstein, I. B. *Molec. cell. Biol.* **9**, 2641–2647 (1989).
- Imber, R. & Fabbro, D. *Cancer Res.* **51**, 632–638 (1991).
- Cepko, C. L., Roberts, B. E. & Mulligan, R. C. *Cell* **37**, 1053–1062 (1984).
- Wigler, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1373–1376 (1979).
- Borner, C., Guadagno, S. N., Hsiao, W.-L. W. & Weinstein, I. B. *Cell Growth Differ.* **1**, 653–660 (1990).
- Jaken, S., Leach, K. & Klaus, T. *J. Cell Biol.* **109**, 697–704 (1989).
- Borner, C., Wyss, R., Regazzi, R., Eppenberger, U. & Fabbro, D. *Int. J. Cancer* **40**, 344–348 (1987).
- Regazzi, R., Fabbro, D., Costa, S. D., Borner, C. & Eppenberger, U. *Int. J. Cancer* **37**, 731–737 (1986).
- O'Brian, C. A., Lawrence, D. S., Kaiser, E. T. & Weinstein, I. B. *Biochem. biophys. Res. Commun.* **124**, 296–302 (1984).
- Costa, S. D., Fabbro, D., Regazzi, R., Kueng, W. & Eppenberger, U. *Biochem. biophys. Res. Commun.* **133**, 814–822 (1985).

ACKNOWLEDGEMENTS. We thank N. Mazurek for the full-length mutant (UV25) and wild-type murine PKCα cDNAs, S. Hyatt and S. Jaken for PKC immunofluorescence analysis and R. S. Krauss, S. N. Guadagno and K. O'Driscoll for reading the manuscript. This work was supported by the Swiss National Science Foundation (I.F. and R.I.), by the National Cancer Institute grant and the Markey Charitable Trust (I.B.W.). C.B. is the recipient of a postdoctoral fellowship supported by the Swiss National Science Foundation.

* To whom correspondence should be addressed.

METHODS. The *S. pombe* conditional lethal mutant *wee1-50 leu1-32 ura4-D18 (adh-cdc25⁺ ura4⁻)::cdc25^h* was transfected with a primary human foreskin fibroblast cDNA library together with the pAL19 library transducing

WEElHu	82	TEFHELEKIGSGEFGSVFKCVKRL--DGCITAIKRSKPLAGSV--DEQNALREYVAHVLQGHSH
weel	563	TRFRNVTLTGSGEFSVEVQVEDPV-EKTLKYAVKKLVKFSGPK-ERNRLLQEVSIQRALKGHDH
mikl	286	IRFQVQVPIHESDFSFVYHVSSINPPTETVAVKMLKNAAKFT-KGERHLQEVSLIQLQACPF
Mos	57	EQVCLLQRLGAGGPGSVKATYRG----VPVAIKQVKNCTKNRLASRSEWAEINLVARL--RHDN
Raf	346	SEVMLSTRIGSGSGFVYKGGWHG----DVAVKILKVDVPTPE-QLOAFRNEVAVLKRT-RHVN

VVRYFSAAWED---DHMLIQNEYCGGSLADAISENYRIMSY-----FKEAELKDLLQVGRGLRY
IVELMDSWEHG---GFLYMVEVLCEGSLDRFLEEQQQLSR-----LDEFVRWKILVEVAVGLQF
VVNLNVNWSYN---DNILFLQLDYCEGQLSLHLSLSELGLLQV-----MDPFRVWKMLFQLTQALNF
IVRVVAASSTRTPAGNSLGTITIMEFGGVNTLHQVIFYAAGHGPEGADGEPHCRTGGQLSLGKLKYSIDLNVNGLLF
ILLFMGYMTK---DNLAIFYTWCNCGSSLYKHLHVOETK-----FQMFLQIDARTAGTGM DY

IHMSLVHMDIKPSNIFIRSTIPNAAASEEGDEDDWANKVMFKIGDLGHVTRISSP-----QVEEGDSRFLA
 IHAKNVVHLDLKPANVMITFEG-----TLKIGDGFQMASVWPVPRG-----MEREGDCEYIA
 IHLLEFVHLDVKPSNVLTISNV-----LITRDGNLKLDFGLATSLFVSSM-----VDLEGRVRYIA
 LHSQSIVHLDLKPANILISEQ-----VCKISDFGCSEKLDLLCFQTPSPVGLGTYTHRA
 LHAKNILHRDKMSKNVIEHLEGL-----TVKIGDGFGLATFVKRSWGSQOV-EQPTGVSVLWMA

NEVLQ--ENYTHLPKADIFALALTVCAAGAEPLRNGDQWHEIROGLRPIQVLSQEFTELLKVMHPDPERR
PEVL--ANHLYDKPADIFSLGITVFEEAANIVLPDNGSQWKLRSQDLSADAPRLSSDTONGSSLTSSSRFPANS
PEIL--ASHNYGKPADVLSLSMIEAATNVLPENGVEQWRLRSGDYSNLPNL--KDLLLSKEKVQINKV
PELL--KGEGVTPKADIYSFAITLWMTQKAPY--GERQHILYAVVAYDLRSLSAVFEDSL-----
PEVMRDNNPFSPQSDVSYSGVILYELMTLGEPLSHYNRRDQIFMGVGRYASPLSRILRYKN-----

PSAMALVKHSL~~LL~~LSASRKS~~AE~~QLRIELNAEKFKNS~~LL~~LQELK/47
II---GGGLDRVVE~~W~~MLSP~~EP~~RNRPTIDQILATDEV~~V~~W~~EM~~R/31
RC---AES~~L~~Q~~CL~~LQRMTHPYVDCRPT~~TD~~QLLAMP~~EM~~IFISEH/17
----PGQLGDV~~IQ~~CRWRPSAA~~OR~~PSAR~~LL~~LV~~DL~~TS~~LS~~KA~~ET~~G*
----CPKAMKRVADCV~~R~~KV~~KE~~ERPL~~FP~~QILSS~~IEL~~LQHS~~LP~~/34

FIG. 2 Amino-acid alignment of *WEE1Hu* protein to other similar kinases. The kinase domain of p49^{WEE1Hu} is aligned to those of *S. pombe wee1*⁺, *S. pombe mik1*⁺, *c-mos* and *c-raf* kinases. Identical matches are shaded. The numbers of additional N- and C-terminal flanking residues lying outside the kinase domain are indicated at the beginning and the end, respectively, of each sequence. A computer-assisted database search found the *S. pombe wee1*⁺ and *c-mos* kinases as those with the highest amino-acid homology.

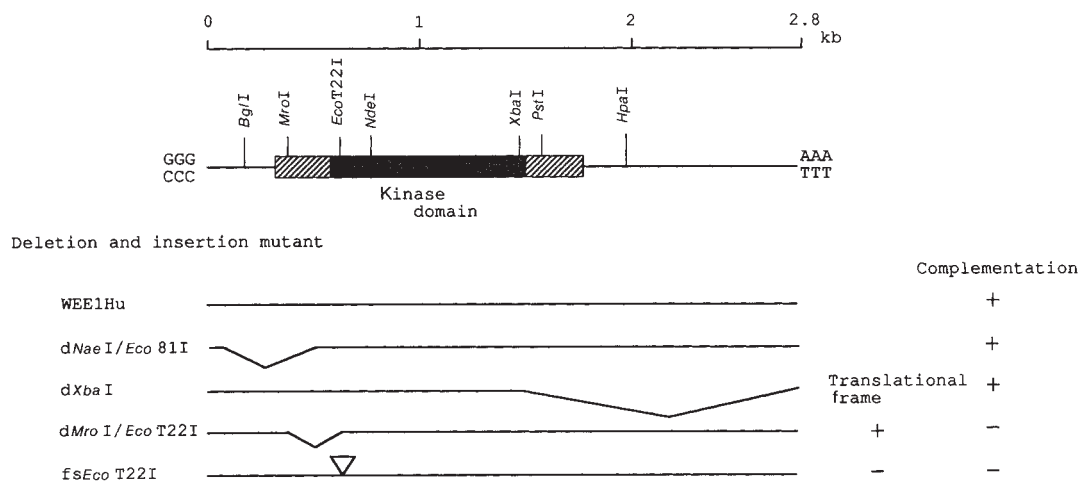


FIG. 3 Deletion analysis of *WEE1Hu*. The restriction map of *WEE1Hu* is shown at the top. The protein-coding region is boxed; the kinase domain is filled and its flanking region is shaded. Disruption of lines and an inverted triangle indicate the positions of deletions and an insertional frameshift, respectively. The putative translation initiation site of the dNaeI-Eco81I mutant is the methionine at nucleotide number 481 (amino-acid number 53). The complementation activity of deletion and insertion mutants is indicated with + (full activity comparable to *WEE1Hu* protein) or - (no

detectable activity).

METHODS. Frame shift and deletion mutants of *WEE1Hu* were constructed by digestion with appropriate restriction enzymes, blunt-ending and religation of a monomeric form of pcD2-*WEE1Hu* that was generated from the originally isolated concatamer. The original as well as mutated *WEE1Hus* were excised with *Bam*HI and inserted into the pcMVL vector, which is a modified pAL7 (ref. 4) vector containing a human cytomegalovirus (CMV) transcriptional unit with the SV40 polyadenylation signal to express inserted cDNA.

of 49,000 (M_r , 49K) and has a typical serine/threonine kinase domain in its central region (underlined).

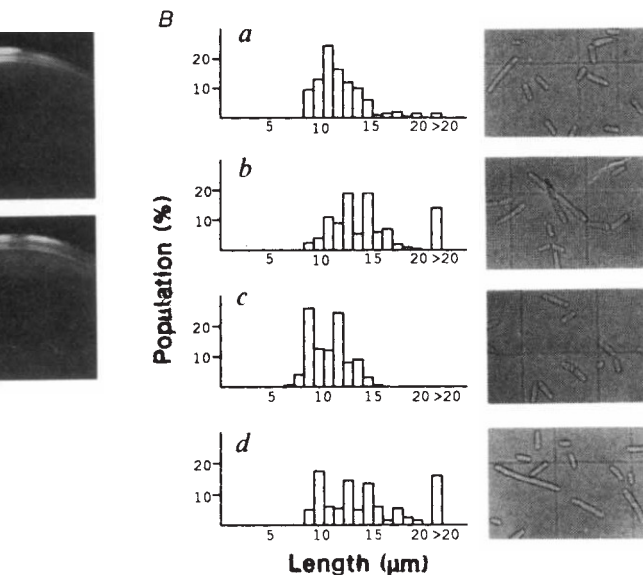
The kinase domain has a significant amino-acid homology with the *S. pombe weel1*⁺, *mik1*⁺ (a functionally redundant sibling of *weel1*⁺) and *c-mos* kinases with the highest similarity to *weel1*⁺ and *mik1*⁺ (29% and 24% amino-acid identities) (Fig. 2). It also has a homology with *c-raf* which is relatively close to the *S. pombe weel1*⁺ kinase in the phylogenetic tree drawn by Hanks *et al.*⁸ The similarity of *WEE1Hu* to *weel1*⁺ and *mik1*⁺ is particularly evident in the C-terminal quarter of the domain where

c-mos and *c-raf* have virtually no homology with this protein. Although *c-mos* and *c-raf* have a significant homology in the core kinase domain, they had no detectable complementation activity even when overexpressed under the human cytomegalovirus promoter, a powerful promoter in *S. pombe*⁹ (data not shown). The N-terminal and C-terminal regions outside the kinase domain of *WEE1Hu* have no similarity to the corresponding regions of the *weel1*⁺ or *mik1*⁺ kinase.

To locate the functional domain of this protein, deletion and frameshift mutants were constructed and tested for their ability

FIG. 4 Effects of *WEE1Hu* expression on yeast cells. A, Complementation assay: a, b, the *S. pombe* strain *adh-cdc25⁺ weel1-50* transfected with *WEE1Hu* (a) or *WEE1Hu* with a frame shift at *EcoT22I* site (b); c, d, the *S. pombe* strain *cdc2-3w weel1-50* transfected with *WEE1Hu* (c) or *WEE1Hu* with a frame shift at *EcoT22I* site (d). B, Effect of *WEE1Hu* overexpression on the length of wild-type yeast cells. a, b, c, d, Cells transfected with pCL-*WEE1Hu*(SV40 promoter) (a), pcMVL-*WEE1Hu* (human cytomegalovirus promoter) (b), pcMVL-*WEE1Hu* with a frame shift at the *EcoT22I* site (c) or the *S. pombe weel1*⁺ gene inserted in the pAL-SK vector (d). The size distribution patterns (left) and the photographs (right) of the cells are shown.

METHODS. A, The *S. pombe* strain *weel1-50 leu1-32 ura4-D18 (adh-ccc25⁺ ura4⁺):cdc25⁺ h⁺* (a, b) or *cdc2-3w weel1-50 leu1-32 h⁻* (c, d) was transfected with pcD2-*WEE1Hu*(a, c) or pcD2-*WEE1Hu* with a frame shift at the *EcoT22I* site (b, d) together with pAL19 (ref. 4). After transfection, cells were spread on minimal medium agar plates¹⁴ and incubated first at 23 °C for 24 h and then at 37 °C (a, b) or at 34 °C (c, d), respectively, for 5 days. B, The *S. pombe* strain *h⁻ leu1-32* was transfected with pCL-*WEE1Hu* (a), pcMVL-*WEE1Hu* (b), pcMVL-*WEE1Hu* with a frame shift at the *EcoT22I* site (c), or the *S. pombe weel1*⁺ gene inserted in the pAL-SK vector (d) and incubated at 30 °C for 3 days on minimal medium agar plates. Colonies were isolated from each



plate, cultured in minimum liquid medium for 10 h and photographed under a microscope. On photographs, 200 cells were randomly chosen from each transformant and measured for cell length. The pCL vector is a modified pAL19 vector⁴ containing an SV40-based transcriptional unit taken from the pcD2 vector. The pcMVL vector is the same as pCL except that the human CMV promoter is used in place of the SV40 early promoter. pAL-SK is pAL7 (ref. 4) containing multiple cloning sites.

to suppress the *wee1* mutation (Fig. 3). As summarized in this figure, only the kinase domain was required for full complementation activity.

The *cdc2* mutant strain *cdc2-3W* phenotypically resembles an overexpressor of *cdc25*⁺, and the *wee1-50 cdc2-3w* double-mutant strain similarly suffers mitotic catastrophe at the nonpermissive temperature¹. Expression of *WEE1Hu* effectively suppressed this lethality too and allowed cells to grow at the nonpermissive temperature (Fig. 4A, c). This suggests that *WEE1Hu* has a function similar to the *wee1*⁺ kinase.

In *S. pombe*, the *wee1*⁺ gene delays the G2-M transition in a dose-dependent manner, resulting in the formation of elongated cells. When *WEE1Hu* was expressed in wild-type *S. pombe* under the control of the simian virus 40 (SV40) promoter, a small population of the cells elongated detectably (more than 17 μ m) (Fig. 4B,a). A remarkable effect was observed when *WEE1Hu* was expressed under the stronger human cytomegalovirus promoter⁹ (there was 5–10 times more *WEE1Hu* mRNA with this promoter): highly elongated cells (>20 μ m) were generated in a 15% population. These elongated cells were similar to, or often longer than, those induced by the *wee1*⁺ gene expressed with a multicopy plasmid (Fig. 4B,b and B,d). The longest cells reached more than 60 μ m (4–6 times as long as wild-type cells). Alternatively, a frameshift mutation in the kinase domain of *WEE1Hu* completely abolished the formation of the highly elongated cells, yielding cells with an average length of 10.5 ± 1.9 μ m, a typical value for wild-type *S. pombe* (Fig. 4B,c). The heterogeneity in size distribution was a result of the variability in the copy number of the plasmids propagating in *S. pombe*.

Two untransformed (MRC-5 and HEF) and four transformed (HeLa, HT29 colon adenocarcinoma, T24 bladder carcinoma and SV40-transformed fibroblast) human cell lines were examined for the expression of *WEE1Hu*. All the cells expressed ~3-kb *WEE1Hu* mRNA (data not shown).

These data taken together indicate that a gene structurally as well as functionally similar to the fission yeast *wee1*⁺ gene is present and expressed in human cells. It should be stressed, however, that the gene we identified may not necessarily be the human counterpart of the *S. pombe wee1*⁺ gene. At this point we cannot exclude the possibility that there may be a better homologue yet to be described.

The *cdc25*⁺ gene is a mitotic inducer controlling the G2-M transition of *S. pombe*⁶. Homologues of the *S. pombe cdc25*⁺ gene exist in *Drosophila*¹⁰ and humans¹¹, and their function is conserved among these diverse species. *String* is a *Drosophila* homologue and its mutation causes a G2 arrest of fertilized eggs¹⁰. *String* protein activates maturation promoting factor in *Xenopus* extracts¹², and the *cdc25*⁺ homologues complement the *cdc25* mutation of *S. pombe*¹¹. These data suggest that the system controlling G2 in the cell cycle, which is best characterized in fission yeast, may well be conserved throughout higher eukaryotes. Our finding that a *wee1*⁺-like gene exists in humans seems further to support this view. □

15. Inoue, H., Nojima, H. & Okayama, H. *Gene* **96**, 23–28 (1990).

16. Sanger, F., Nicklen, S., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).

ACKNOWLEDGEMENTS. We thank P. Nurse for providing the *S. pombe* mutant *wee1-50 leu-32 ura4-D18 (adh-cdc25⁺ ura4⁺)::cdc25^h*, and K. Okazaki, B. K. Benton, K. Kume, K. Tanaka, S. Kondoh, Y. Suzuki and Y. Terada for helpful discussions. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan, and the Ryoichi Naito Foundation for Medical Research and Osaka Cancer Society.

Oncogenic germ-line mutations in Sp1 and ATF sites in the human retinoblastoma gene

Toshiyuki Sakai*, Naoko Ohtani*, Terri L. McGee, Paul D. Robbins† & Thaddeus P. Dryja

Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114, USA

† Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA

THE transcription of a eukaryotic gene is a consequence of intricate interactions between members of a set of transcription factors¹. We describe here evidence indicating that at least two distinct DNA-binding factors play an important part in the transcription of the human retinoblastoma gene (*Rb*). One of the factors reacts with a sequence overlapping with a potential Sp1 recognition sequence in the promoter region of the gene, the other with a nearby ATF recognition sequence. We have identified two naturally occurring point mutations in these recognition sequences that cause hereditary retinoblastoma. The nuclear factors do not bind to the mutant sequences. We infer that these nuclear factors are necessary for the expression of the *Rb* gene and the suppression of cancer.

DNA samples from 111 patients with retinoblastoma were screened for mutations of the 5' untranslated region or the promoter region of the *Rb* gene, using the single-strand conformation polymorphism (SSCP) technique². DNA was derived from leukocytes from patients known to have a germ-line mutation (bilateral disease or a positive family history) or from primary tumour fragments. Blood samples with deletions or other gene rearrangements detectable by Southern blotting were not included in this analysis, nor were tumours that were homozygous for such abnormalities. With these inclusion criteria, each sample carries at least one mutation of the *Rb* gene that is beyond the resolution of Southern blotting techniques, and some will carry two (that is, distinct mutations of the two gene homologues in tumours that retain heterozygosity). In all, 125 mutant alleles were screened.

Two cases showed a sequence variation within the region extending 327 bases upstream of the initiation codon. In both instances, the variation was present in leukocyte DNA from all affected members or obligate carriers of families with hereditary retinoblastoma that we examined (Fig. 1). DNA sequence analysis showed that in family number 172 the variant band was due to a G-to-T transversion 189 base pairs (bp) upstream of the initiating methionine codon, and that in family 229 it was due to a G-to-A transition 9 bp further upstream. The penetrance of these mutations appeared to be low: both carriers in family 172 had only unilateral retinoblastoma, and there were at least three obligate carriers who had no retinoblastoma in family 229 (members II-3, III-3 and III-8). A concurrent analysis of the majority of the coding region of the *Rb* gene also using the SSCP technique has revealed no other abnormalities in the DNA sequence in these two cases (D. Yandell, personal communica-

* Present address: Department of Preventive Medicine, Kyoto Prefectural University of Medicine, Kawaramachi-Hirokoji, Kamikyo-ku, Kyoto 602, Japan.

Received 22 April; accepted 18 July 1991.

1. Russell, P. & Nurse, P. *Cell* **49**, 559–567 (1987).
2. Featherstone, C. & Russell, P. *Nature* **349**, 808–811 (1991).
3. Lee, M. & Nurse, P. *Nature* **327**, 31–35 (1987).
4. Okazaki, K., Okazaki, N., Kume, K., Jinno, S. & Okayama, H. *Nucleic Acids Res.* **18**, 6485–6489 (1990).
5. Lundgren, K. et al. *Cell* **64**, 1111–1122 (1991).
6. Russell, P. & Nurse, P. *Cell* **45**, 145–153 (1986).
7. Chen, C. & Okayama, H. *Molec. cell. Biol.* **7**, 2745–2752 (1987).
8. Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 24–32 (1988).
9. Toyama, R. & Okayama, H. *FEBS Lett.* **268**, 217–221 (1990).
10. Edgar, B. A. & O'Farrell, P. H. *Cell* **57**, 177–187 (1989).
11. Sadhu, K., Reed, S. I., Richardson, H. & Russell, P. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5139–5143 (1990).
12. Kumagai, A. & Dunphy, W. G. *Cell* **64**, 903–914 (1991).
13. Gutz, H., Heslot, H., Leupold, U. & Loprieno, N. in *Handbook of Genetics* Vol. 1 (ed. King, R. D.) 395–446 (Plenum, New York, 1974).
14. Rothstein, R. in *DNA Cloning—A Practical Approach* Vol. 2 (ed. Glover, D. M.) 45–65 (IRL, Oxford, 1985).

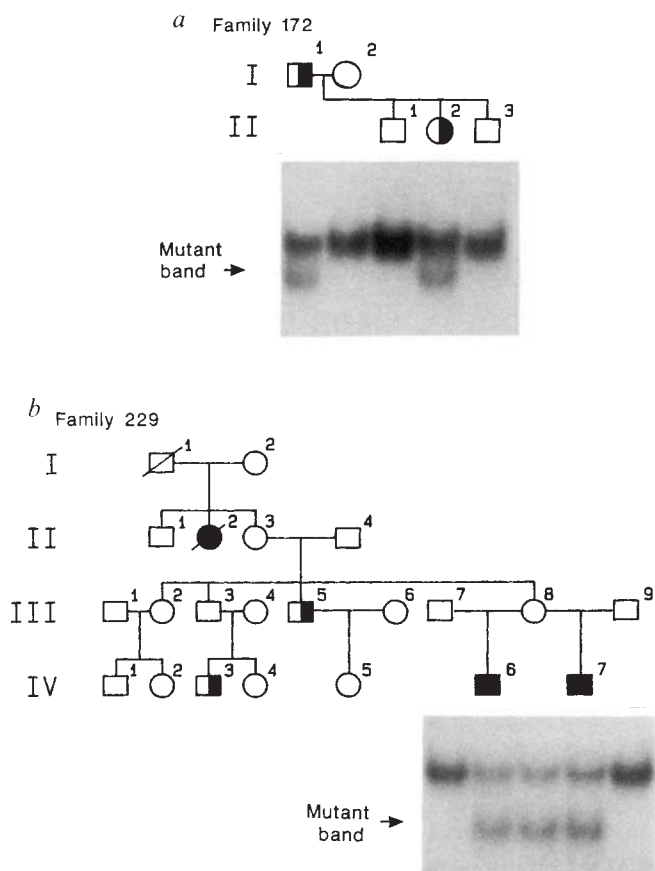


FIG. 1 SSCP analysis of members of families 172 and 229. *a*, Individuals I-1 and II-2 of family 172 developed unilateral retinoblastoma and are heterozygous for a mutant band. *b*, Only a portion of family 229 provided blood samples for analysis. Members III-8, IV-6 and IV-7 carry a mutation; II-3, an obligate carrier, developed breast cancer. Filled symbols denote bilateral retinoblastoma, half-filled symbols unilateral retinoblastoma. METHODS. Approximately 100 ng leucocyte DNA was used as a template for the polymerase chain reaction (PCR) with the following primers: sense, 5'-CTGGACCCACGCCAGGTTTC-3'; antisense, 5'-ATTGGTACCCGACTCCCGTTACAAAAT-3'. The target sequence was 239 bp and extended from -327 to -89 with respect to the first base of the initiation codon⁹. PCR was performed according to a modification of published methods². To increase the sensitivity of the SSCP technique, the amplified DNA was cleaved with *Apal* before it was heat-denatured and loaded onto a non-denaturing 6% polyacrylamide gel.

tion). We found neither of these mutations among 111 unrelated, normal 'control' individuals.

These two mutations are within the consensus recognition sequences of two distinct types of transcription factors. The mutation in family 172 is within a sequence homologous to the consensus sequence for ATF, a member of a large family of related transcription factors that includes CREB (ref. 3). The mutation in family 229 is within a sequence similar to that recognized by Sp1, a mammalian transcription factor with a zinc-finger DNA-binding domain¹.

We next used a gel-shift assay to determine whether nuclear factors bind to these regions and if the identified mutations affect this binding. A nuclear extract from CV-1 cells (a monkey kidney cell line) was incubated with matched sets of oligomers based on the normal and mutant sequences from this region of the *Rb* gene. The nuclear extract retarded the migration through a polyacrylamide gel of the normal ATF sequence but not the one with the mutation found in family 172. Furthermore, the labelled complex was diluted by the addition of unlabelled wild-type sequence but not with unlabelled mutant sequence (Fig. 2a). In addition, the binding of this factor was not affected by the nearby mutation found in family 229 in a nearby consensus Sp1-binding sequence. We concluded that a nuclear factor(s) forms a complex with the wild-type but not the mutant sequence, and that this factor may be a member of the ATF/CREB family.

With regard to the possible Sp1 site, the CV-1 nuclear extract reacted with the wild-type sequence from the *Rb* gene but not to the mutant sequence derived from family 229 (Fig. 2b). In an attempt to characterize this factor, we synthesized a consensus Sp1-binding site that strongly reacts with the Sp1 protein (GGGGCGGGGC) (ref. 4) and which is different from the Sp1-like sequence found in this region of the *Rb* gene (CGCC-GCGGGCGG). This unlabelled consensus Sp1 recognition sequence did not compete with the *Rb* gene sequence for binding to the CV-1 nuclear factor (Fig. 2b). To study this further, we

repeated the assay using nuclear extracts from F9 cells, in which binding of the Sp1 protein to its consensus recognition sequence is easily detected. As shown in Fig. 2c, the labelled Sp1 consensus sequence reacts with a nuclear factor from F9 cells. We have demonstrated that the factor interacting with this consensus sequence is indeed Sp1, using Sp1 purified both from HeLa cells and from bacteria containing an Sp1 expression plasmid (data not shown). The detection of this complex, denoted as complex A in Fig. 2c, is decreased by the addition of competing cold DNA with the normal but not the mutant sequence from the *Rb* gene. This suggests that the Sp1 protein can bind to the *Rb*-Sp1 site, but has a lower affinity for the wild-type sequence from the *Rb* gene than for the Sp1 consensus sequence. It thus appears that a factor producing a different gel shift (complex B in Fig. 2c) binds with higher affinity to the wild-type sequence from the *Rb* gene. Detection of this complex is decreased when unlabelled wild-type *Rb* gene sequence is added, but not when the mutant sequence from family 229 or the consensus Sp1 sequence is added. Our interpretation of these results is that wild-type sequence from the *Rb* gene is recognized by a nuclear protein, distinct from Sp1, which binds to a sequence that overlaps with the Sp1 recognition sequence. We term this factor retinoblastoma binding factor 1, or RBF-1.

It is likely that the nuclear factors described here are important for the expression of the *Rb* gene as the mutations in the binding sites cosegregate with predisposition to retinoblastoma in the respective families. To substantiate this inference, we constructed plasmids with either the normal or mutant sequences inserted upstream of a reporter gene, luciferase. In addition, we deleted portions of the normal promoter sequence in the same vector. We found that the expression of the reporter gene in CV-1 cells depended on sequences between 206 and 185 bases upstream of the initiation methionine codon (Fig. 3). Furthermore, the reporter gene activity was substantially decreased by either of the point mutations.

Our findings offer a rudimentary understanding of the tran-

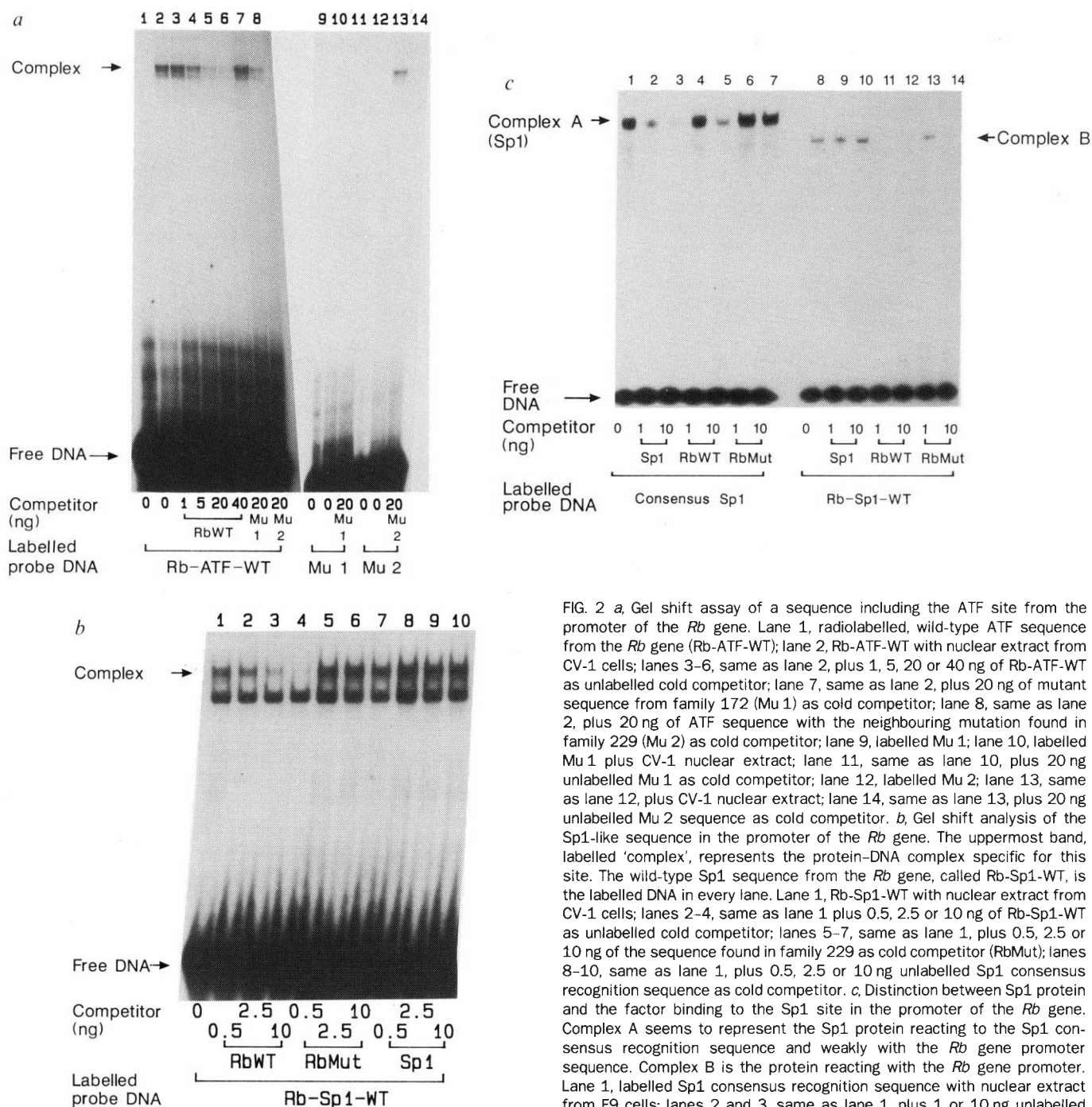
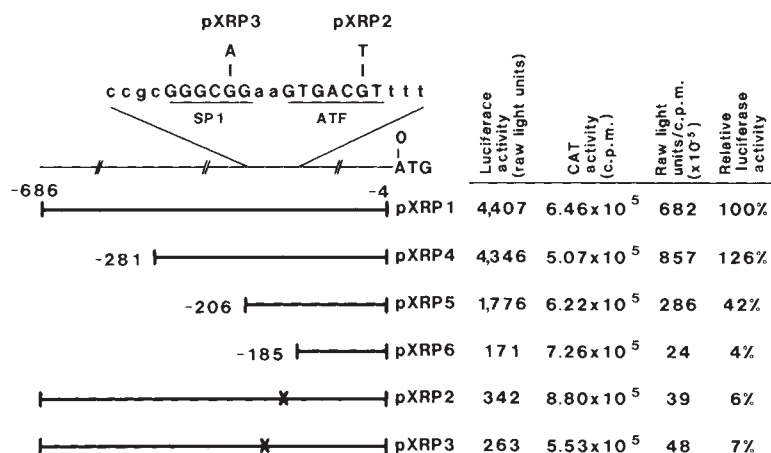


FIG. 2 *a*, Gel shift assay of a sequence including the ATF site from the promoter of the *Rb* gene. Lane 1, radiolabelled, wild-type ATF sequence from the *Rb* gene (Rb-ATF-WT); lane 2, Rb-ATF-WT with nuclear extract from CV-1 cells; lanes 3–6, same as lane 2, plus 1, 5, 20 or 40 ng of Rb-ATF-WT as unlabelled cold competitor; lane 7, same as lane 2, plus 20 ng of mutant sequence from family 172 (Mu 1) as cold competitor; lane 8, same as lane 2, plus 20 ng of ATF sequence with the neighbouring mutation found in family 229 (Mu 2) as cold competitor; lane 9, labelled Mu 1; lane 10, labelled Mu 1 plus CV-1 nuclear extract; lane 11, same as lane 10, plus 20 ng unlabelled Mu 1 as cold competitor; lane 12, labelled Mu 2; lane 13, same as lane 12, plus CV-1 nuclear extract; lane 14, same as lane 13, plus 20 ng unlabelled Mu 2 sequence as cold competitor. *b*, Gel shift analysis of the Sp1-like sequence in the promoter of the *Rb* gene. The uppermost band, labelled 'complex', represents the protein-DNA complex specific for this site. The wild-type Sp1 sequence from the *Rb* gene, called Rb-Sp1-WT, is the labelled DNA in every lane. Lane 1, Rb-Sp1-WT with nuclear extract from CV-1 cells; lanes 2–4, same as lane 1 plus 0.5, 2.5 or 10 ng of Rb-Sp1-WT as unlabelled cold competitor; lanes 5–7, same as lane 1, plus 0.5, 2.5 or 10 ng of the sequence found in family 229 as cold competitor (RbMut); lanes 8–10, same as lane 1, plus 0.5, 2.5 or 10 ng unlabelled Sp1 consensus recognition sequence as cold competitor. *c*, Distinction between Sp1 protein and the factor binding to the Sp1 site in the promoter of the *Rb* gene. Complex A seems to represent the Sp1 protein reacting to the Sp1 consensus recognition sequence and weakly with the *Rb* gene promoter sequence. Complex B is the protein reacting with the *Rb* gene promoter. Lane 1, labelled Sp1 consensus recognition sequence with nuclear extract from F9 cells; lanes 2 and 3, same as lane 1, plus 1 or 10 ng unlabelled Sp1 consensus recognition sequence as cold competitor; lanes 4 and 5, same as lane 1, plus 1 or 10 ng of unlabelled Rb-Sp1-WT as cold competitor; lanes 6 and 7, same as lane 1, plus 1 or 10 ng unlabelled RbMut as cold competitor; lane 8, labelled Rb-Sp1-WT with nuclear extract from F9 cells; lanes 9 and 10, same as lane 8, plus 1 or 10 ng of unlabelled Sp1 consensus recognition sequence; lanes 11 and 12, same as lane 8, plus 1 or 10 ng unlabelled Rb-Sp1-WT as cold competitor; lanes 13 and 14, same as lane 8, plus 1 or 10 ng unlabelled RbMut as cold competitor. **METHODS.** Nuclear extracts were prepared as described¹⁰; the gel shift assay was done according to a modification of the procedure described in ref. 11. Radiolabelled double-stranded DNA was made by annealing complementary oligomers with 5' overhangs and then filling in the recessed 3' ends with ³²P-labelled nucleotides using the Klenow fragment of DNA polymerase. All assayed sequences have flanking partial *Hind*III and *Sal*I recognition sequences not present in the *Rb* gene. The resulting DNA fragments had the following sequences (complete sense strands): Rb-ATF-WT, AGCTCGGAAGTGACGTTTTCCCTCGA; Mu 1, AGCTCGGAAGTGACTTTTTCCCTCGA; Mu 2, AGCTCAGAAGTGACGTTTTCCCTCGA; Rb-Sp1-WT, AGCTGCGCGGGCGGAAGTTTCGA; RbMut, AGCTGCGCGGGCGGACAGATTTCGA; Sp1 consensus recognition sequence: AGCTGGGGCGGGGCTCGA.

scriptional regulation of the *Rb* gene, a human tumour-suppressor gene. We demonstrate the existence of binding sites for two nuclear factors essential for its transcription. On the basis of its recognition sequence, one of the factors may be a member of the ATF family of transcription factors. The other, RBF-1, reacts with a sequence that overlaps with an Sp1-binding site and that is weakly recognized by Sp1. It is not clear from our data whether the decrease in transcription caused by a point mutation in this sequence is due to interference of the binding of RBF-1 and/or Sp1. The contiguity of these two DNA binding sites makes us wonder whether the two factors interact with each other to allow transcription of the *Rb* gene. To our knowledge, this is the first description of naturally occurring mutations in such binding sites as a cause of human disease. Our results are in agreement with those of Hong *et al.*⁵, who found that the sequence extending from 196 to 249 bases upstream of the initiation codon is essential for transcription. Additional

FIG. 3 Relative ability of sequences from the 5' end of the *Rb* gene to drive the expression of a reporter gene, luciferase, in CV-1 cells. The numbers at the ends of the horizontal bars indicate the extent of each tested DNA fragment relative to the first base of the initiation codon. Plasmids pXRP2 and pXRP3 are identical to pXRP1 except for the single base substitutions found in families 172 and 229, respectively.

METHODS. All promoter sequences used in this study were amplified from human genomic DNA using primers with a synthetic *Bam*HI and *Xho*I sites at the 5' and 3' ends, respectively. Amplified fragments were digested with *Bam*HI and *Xho*I and subcloned into the luciferase vector pXP2 (ref. 12). Clones with sequence abnormalities due to cloning artefacts were discarded. Plasmid DNA was transfected into 10^6 CV-1 cells using DEAE-dextran. In each case, $2 \mu\text{g}$ of the tested plasmid was cotransfected with $2 \mu\text{g}$ pSV2CAT, a plasmid with the chloramphenicol acetyl transferase gene (CAT) which is constitutively expressed in CV-1 cells and used here as an internal control. Two days after transfection, luciferase activity and CAT activity were measured^{13,14}. Relative luciferase activity represents the measured luciferase activity divided by the measured CAT activity, both in arbitrary units, and normalized so that pXRP1 gives 100%. Values are the mean based on one experimental run of three transfections for each plasmid. Two additional experimental runs gave similar results.



evidence that this region is important comes from the report of a prostate carcinoma with a deletion of 103 bases that overlaps this region⁶.

Future studies of the factors binding to these sites should provide insight into their role in initiating transcription of the *Rb* gene and thereby suppressing cancer. Examination of the set of genes regulated by these factors might show that it includes other tumour-suppressor genes or proto-oncogenes. We note that the promoter region of the *c-fos* proto-oncogene carries a functional ATF site with remarkable similarity to the one we have found in the *Rb* gene⁷.

The families with these mutations have 'low-penetrance' retinoblastoma, namely frequent occurrence of asymptomatic or unilaterally affected carriers⁸. This may be related to the fact that the mutations described here result in a quantitative decrease in the expression of the *Rb* gene, rather than in its complete inactivity. Perhaps the low level of expression is near the boundary that distinguishes benign from malignant cells. Homozygosity for these weak mutations may leave a sensitive cell on the 'benign' side of the distinguishing boundary, thereby excluding a common route by which the homologous normal allele is eliminated before oncogenesis. To develop into a retinoblastoma, it may be necessary for a retinal cell carrying one of these mutations to acquire a second, complete loss-of-function mutation of the homologous normal allele. As such 'second hits' are much less frequent than chromosomal homozygosity (30% of cases versus 70%), the proportion of affected individuals would be low. □

Received 30 April; accepted 12 July 1991.

- Mitchell, P. J. & Tjian, R. *Science* **245**, 371–245 (1989).
- Orita, M., Suzuki, Y., Sekiya, T. & Hayashi, K. *Genomics* **5**, 874–879 (1989).
- Hai, T., Liu, F., Coukos, W. J. & Green, M. R. *Genes Dev.* **3**, 2083–2090 (1989).
- Kadonaga, J. T., Courey, A. J., Ladika, J. & Tjian, R. *Science* **242**, 1566–1570 (1988).
- Hong, F. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5502–5506 (1989).
- Bookstein, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7762–7766 (1990).
- Sassone-Corsi, P., Ransone, L. J. & Verma, I. M. *Oncogene* **5**, 427–431 (1990).
- Macklin, M. T. *Am. J. hum. Genet.* **12**, 1–43 (1960).
- McGe, T. L., Yandell, D. W. & Dryja, T. P. *Gene* **80**, 119–128 (1989).
- Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475–1489 (1983).
- Distel, R. J., Ro, H.-S., Rosen, B. S., Groves, D. L. & Spiegelman, B. M. *Cell* **49**, 835–844 (1987).
- Nordeen, S. K. *BioTechniques* **6**, 454–457 (1988).
- Brasier, A. R., Tate, J. E. & Habener, J. F. *BioTechniques* **7**, 1116–1122 (1989).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning* **3**, 16–63 (Cold Spring Harbor Laboratory Press, New York, 1989).

ACKNOWLEDGEMENTS. We thank D. Yandell and J. Toguchida for discussion, J. Rapaport and U. Onwuta for technical assistance, and T. Nagaya, A. Brasier, T. Kay, T. Nikaído, T. Meyer and J. Horowitz for help and advice. This work was supported in part by the NIH.

Putative receptor for the cytoplasmic inactivation gate in the *Shaker* K⁺ channel

Ehud Y. Isacoff, Yuh Nung Jan & Lily Yeh Jan

Howard Hughes Medical Institute and the Departments of Physiology and Biochemistry, University of California, San Francisco, California 94143, USA

INACTIVATION of ion channels is important in the control of membrane excitability. For example, delayed-rectifier K⁺ channels, which regulate action potential repolarization, are inactivated only slowly, whereas A-type K⁺ channels, which affect action potential duration and firing frequency, have both fast and slow inactivation¹. Fast inactivation of Na⁺ and K⁺ channels may result from the blocking of the permeation pathway by a positively charged cytoplasmic gate² such as the one encoded by the first 20 amino acids of the *Shaker* B (ShB) K⁺ channel^{3,4}. We report here that mutation of five highly conserved residues between the proposed membrane-spanning segments S4 and S5 (also termed H4)⁵ of ShB affects the stability of the inactivated state and alters channel conductance. One such mutation stabilizes the inactivated state of ShB as well as the inactivated state induced in the delayed-rectifier type K⁺ channel drk1 (ref. 6) by the cytoplasmic application of the ShB N-terminal peptide. The S4–S5 loop, therefore, probably forms part of a receptor for the inactivation gate and lies near the channel's permeation pathway.

As neutralizations of basic residues in the N terminus of ShB reduce fast inactivation³, we asked if acidic residues interacting with the N terminus could be identified and, if so, whether they lie near the permeation pathway. Among several glutamine substitutions of highly conserved acidic residues modelled to be on the cytoplasmic side of the membrane (E200, E208, E218 and E395) only one, E395Q, affected inactivation, greatly reducing the fast component of macroscopic inactivation (at 11 °C; Fig. 1b). The slower component of macroscopic inactivation was also affected (Table 1). Previous studies^{7,8} have shown that, at highly depolarized membrane potentials, the fast component of macroscopic inactivation is determined largely by the voltage-independent rate of transition of a single channel from the open

TABLE 1 Properties of channels with point mutations in the S4-S5 loop

Mutant	Activation voltage dependence		Reversal potential			N	A2/(A1 + A2)	Onset of inactivation		N	Recovery from inactivation		Single-channel conductance		N
	V _{0.5} (mV)	B (mV)	V _r (mV)					TAU1 (ms)	TAU2 (ms)		TAU (ms)	N	+100 mV (pS)	-100 mV (pS)	
			20 mM [K ⁺] _o	40 mM [K ⁺] _o	89 mM [K ⁺] _o										
L385F	29 ± 3	18 ± 1	-28 ± 8	-16 ± 6	nd	3	0.24 ± 0.07	38 ± 11	869 ± 197	6	67 ± 11	5	nd	nd	—
L385V	12 ± 1	21 ± 3	-39 ± 2	-24 ± 1	ND	3	0.12 ± 0.02	28 ± 9	1427 ± 514	5	130 ± 12	4	13.2 ± 0.3	23.6 ± 0.5	5
L385A	-3 ± 3	20 ± 3	-26 ± 3	-12 ± 2	4 ± 1	4	0.06 ± 0.02	26 ± 3	346 ± 225	10	>500	5	4.8 ± 0.9	12.2 ± 1.2	2
T388S	-1 ± 4	19 ± 2	-31 ± 3	-17 ± 1	0 ± 2	4	0.87 ± 0.04	34 ± 3	6887 ± 3118	4	ND	—	9.8 ± 1.4	18.7 ± 1.7	3
T388A	5 ± 5	21 ± 1	-30 ± 1	-16 ± 1	0 ± 2	4	0.76 ± 0.04	29 ± 5	5658 ± 1415	7	40 ± 6	4	8.8 ± 0.8	15.0 ± 0.9	8
L389A*	-9	19	ND	ND	ND	1	0.42 ± 0.04	16 ± 2	1748 ± 303	7	62 ± 3	5	10.3 ± 0.3	26.6 ± 1.1	3
K390Q*	-5 ± 4	21 ± 1	ND	ND	ND	3	0.34 ± 0.02	20 ± 2	1168 ± 158	4	53 ± 5	5	11.1 ± 0.4	22.6 ± 2.0	3
S392C	-5 ± 2	18 ± 1	-30 ± 5	-18 ± 5	ND	4	0.92 ± 0.09	21 ± 7	1783 ± 270	6	ND	—	5.4 ± 0.3	13.3 ± 0.7	5
S392A	-3 ± 4	14 ± 3	-34 ± 3	-18 ± 3	-2 ± 2	3	0.57 ± 0.10	22 ± 5	2071 ± 471	6	ND	—	7.0 ± 1.2	15.9 ± 2.4	5
R394Q*	-10 ± 4	19 ± 2	ND	ND	ND	3	0.27 ± 0.08	17 ± 1	2054 ± 344	5	48 ± 7	4	11.0 ± 0.7	23.1 ± 1.3	3
E395D	nd	nd	ND	ND	ND	—	1.00	—	4557 ± 771	3	ND	—	8.6 ± 1.3	†	5
E395Q	6 ± 3	20 ± 1	-34 ± 3	-19 ± 3	-4 ± 1	5	1.00	—	7343 ± 1037	3	ND	—	†	†	—
L396A	-6 ± 4	21 ± 1	-28 ± 1	-14 ± 3	ND	3	0.88 ± 0.02	18 ± 5	3738 ± 321	3	22 ± 1	2	8.2 ± 0.5	16.8 ± 1.2	5
G397A*	3	21	ND	ND	ND	1	0.30 ± 0.08	21 ± 1	1310 ± 91	6	47 ± 5	6	10.1 ± 0.9	18.4 ± 0.8	2
ShB‡	-7 ± 4	16 ± 3	-34 ± 3	-20 ± 1	-2 ± 2	6/5	0.24 ± 0.06	16 ± 2	1323 ± 274	9	55 ± 7	4	11.7 ± 0.9	24.2 ± 2.0	6

Voltage dependence of activation was determined from normalized conductance-voltage curves. Data were obtained in MBSH (11 ± 1 °C; see legend to Fig. 1) by two-electrode voltage clamping using 0.5-s long pulses from a holding potential of -100 mV to step potentials ranging from -70 to +80 mV in 10 mV increments. Conductance was measured as the peak amplitude of the leak-subtracted current divided by the driving force using the extrapolated reversal potential (V_r). Data were fit with the Boltzmann equation ($G/G_{\max} = 1/(1 + \exp((V_{0.5} - V)/B))$), where V is the step potential. The midpoint potential, $V_{0.5}$, and the slope parameter, B , were determined separately for each oocyte. Conductance-voltage values were not determined for E395D owing to its low amount of expression. Because of the small size and fast decay of the tail current in normal (1 mM) external K^+ , V_r was first determined at 20, 40 and 89 mM external K^+ and a linear regression on the semilogarithmic plot of V_r versus external K^+ concentration ($[K^+]_o$) was used to extrapolate to V_r values at 1 mM K^+ . Macroscopic inactivation parameters were measured using PCLAMP's least squares minimization for two exponential fits to two-electrode voltage clamp currents (11 ± 1 °C) evoked by 0.4, 2 and 10-s-long depolarizing pulses to +40 mV (equal to $G_{\max}$) from a holding potential of -100 mV (except in the case of L385F for which steps to +80 mV were used to reach the voltage shifted $G_{\max}$ (see ref. 14)). Chord conductance measurements at indicated potentials were calculated from data similar to those presented in Fig. 3, using the reversal potential of 0 mV in the symmetric K^+ solutions. Chord conductances (in pS) at +100 mV/-100 mV for mutants not included in this table were: 11.6 ± 0.2/23.7 ± 1.2 ($N=3$) for ShB(Δ17-25); 7.1 ± 1.0/13.8 ± 1.0 ($N=3$) for the double mutant ShB(Δ17-25)/T388A; 5.5 ± 0.4/12.4 ± 0.5 ($N=5$) for the double mutant ShB(Δ17-25)/S392C; and 8.5/ND ($N=1$) for the double mutant ShB(Δ17-25)/E395D. Values are mean ± s.d. for N oocytes.

* Conductance-voltage relations were determined assuming $V_r = -100$ mV.

† Conductance could not be measured owing to the shortness of both outward and inward openings for E395Q and inward openings for E395D.

‡ V_r values for ShB ($N=5$) were determined alongside those of the mutants. The $V_{0.5}$ and B values for ShB are taken from previous determinations ($N=6$) made under identical conditions by G. Lopez (manuscript in preparation). Inactivation kinetics for ShB have already been reported²⁰.

ND Not determined.

state to an inactivated state whereas the slower component is due to reopenings from this fast inactivated state. The decay of this slow component is associated with the transition into a distinct 'slow' inactivated state.

E395 is in a loop joining the two proposed membrane spanning segments S4 and S5. In this S4-S5 loop (Fig. 1a), mutations of two polar residues, a threonine (T388, to serine or alanine) and a serine (S392, to cysteine or alanine), and of a nonpolar leucine (L396, to alanine) all decreased the relative amplitude of the fast component of macroscopic inactivation as did mutations of the glutamate (E395, to aspartate or glutamine) (Fig. 1b, Table 1). In contrast, a valine for leucine (L385) mutation increased the relative amplitude of the fast component of inactivation and an alanine substitution had a similar but more extreme effect. A phenylalanine substitution at L385 did not alter the relative amplitudes of the inactivation components but accelerated slow inactivation as did the alanine substitution (Fig. 1b, Table 1). Mutations at other positions in the S4-S5 loop (L389, K390, R394 and G397) had little effect on inactivation (Fig. 1b, Table 1).

The effects of the mutations on the extent of fast inactivation correlated with changes in the rate of recovery from fast inactivation at hyperpolarized potentials (at 11 °C; Fig. 1c), as well as with changes in the rate of reopening from the fast inactivated state of single channels at depolarized potentials (at 21-24 °C; Fig. 2a, b). Furthermore, for two cases in which the mutants expressed sufficiently large currents to allow patch recording of macroscopic current (at 21-24 °C) there was a quantitative agreement between the effect of the mutation on macroscopic inactivation and on the rate of reopening of single channels (Fig. 2c), indicating that these mutations primarily altered the life time, that is the stability of the fast-inactivated state.

Two other aspects of channel function, the voltage dependence of activation (the conductance-voltage curve) and the dependence of currents on external K^+ (reversal potential), were only mildly, or not at all, affected by most of the mutations

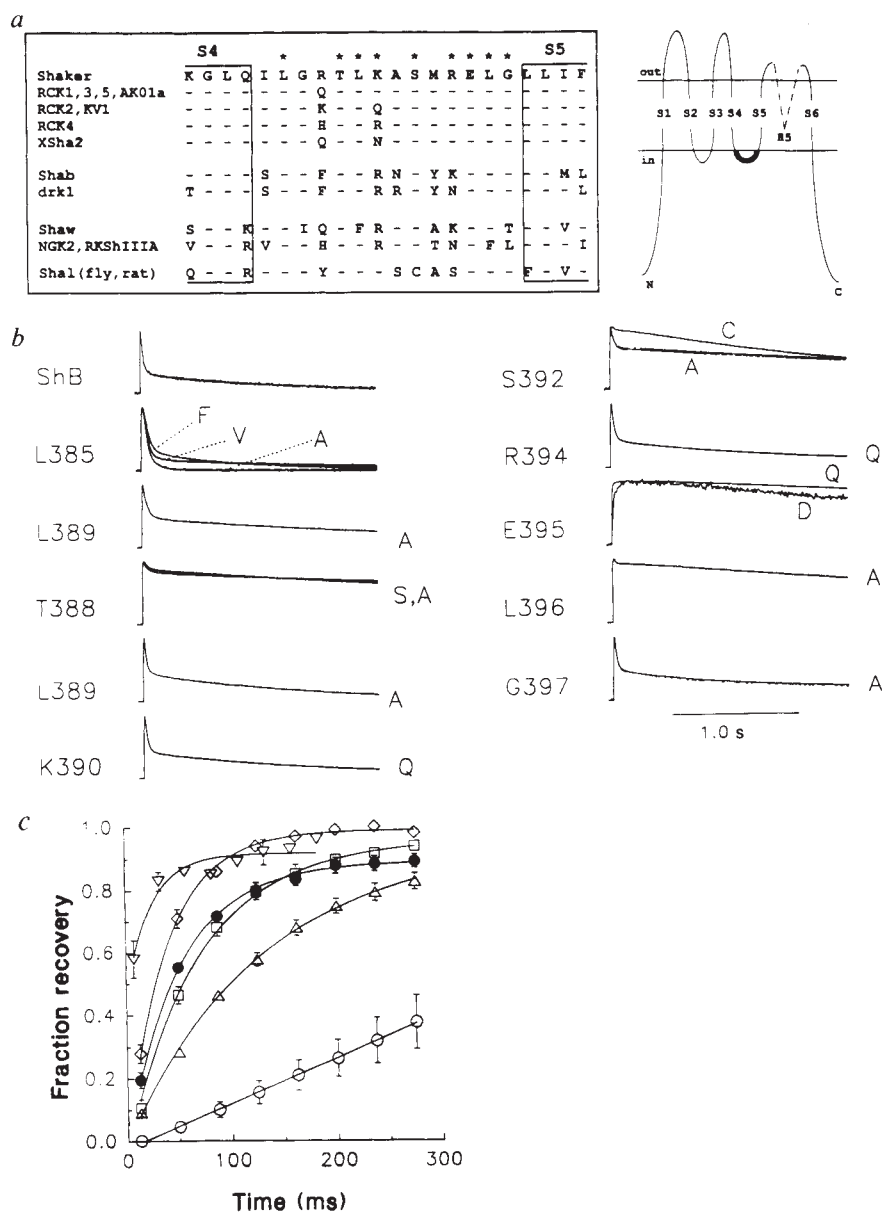
described here (Table 1). The notable exceptions were shifts in the conductance-voltage curve in two mutations at L385 (L385V and L385F, Table 1), which is close to the S4 sequence (Fig. 1a) that has been proposed to be the channel's voltage sensor⁹⁻¹⁴. The preservation in the S4-S5 loop mutants of other normal channel functions indicates that the effects on inactivation were not likely to be due to global disruptions of channel structure.

Among the five residues whose mutation affected channel inactivation, L385 in the wild-type ShB K^+ channel seems to have the special role of destabilizing the fast-inactivated state, allowing for multiple openings during a sustained depolarization and rapid recovery during brief repolarization. Replacing this residue with residues that have smaller side chains stabilized the fast-inactivated state, perhaps by allowing tighter binding of the cytoplasmic inactivation gate to its receptor. This possibility was tested in drk1, a delayed-rectifier type K^+ channel with no detectable fast inactivation⁶, by mutating the leucine at the equivalent position (L316).

Cytoplasmic application of the ShB N-terminal peptide introduced a fast component of macroscopic inactivation to drk1 (Fig. 2d, left), even though drk1 belongs to a subfamily of K^+ channels different from *Shaker* and RCK1 (refs 6, 15, 16), which have been found to interact with the ShB peptide⁴. The peptide introduced a new long-lived closed state to drk1 that interrupted normal channel bursting (Fig. 2d, middle and bottom left). The L316A mutation of drk1 prolonged the inactivated closed times induced by the peptide (Fig. 2d, right). The mutation also increased the rate of slow inactivation of the channel (Fig. 2d, top right) as did the equivalent mutation, L385A, in ShB (Table 1). Therefore, as with *Shaker*, this alanine for leucine substitution stabilized the inactivated state caused by the interaction between the drk1 channel and the ShB peptide. Taken together, the effects of the ShB and drk1 mutants (Fig. 2d, Table 1) are consistent with the model that the S4-S5 loop is a receptor for the fast inactivation gate and suggest that the processes of fast and slow inactivation interact.

FIG. 1 Mutations in the S4-S5 loop affect channel inactivation. **a**, Alignment of S4-S5 loop sequences of 16 cloned K⁺ channels (broken down into the *Drosophila* Shaker, Shab, Shaw and Shal subfamilies and their homologues in other species) reveals a high degree of amino-acid identity (represented as dashes in sequences other than *Shaker* protein) at several positions (single-letter amino-acid code). The glutamate (E) in the *Shaker* protein sequence is E395. Residues marked with asterisks were mutated in this study. Clones in the same row are identical over region shown. References for the clones are provided in ref. 16, except for AK01a (ref. 26), XSh2 (ref. 27), SKShIIIA (ref. 28) and rat Shal (ref. 29). Insert indicates location of the S4-S5 loop in a transmembrane topology model of K⁺ channels³⁰. H5 (broken line) has been modelled to dip partway into the membrane as two short helices connected by random coil³⁰ or as a beta sheet with a hairpin turn²⁴. **b**, Normalized macroscopic two-electrode voltage clamp records show that mutations at a subset of positions in the S4-S5 loop (L385, T388, S392, E395 and L396) alter the relative amplitude of the fast component of inactivation. (Although the fast component of inactivation is absent in E395Q in two-electrode voltage-clamp recordings (11 ± 1 °C) a fast component was evident in macroscopic patch recordings (21–24 °C), probably due to a greater synchronization of inactivation transitions at the higher temperature.) **c**, The rate of recovery from inactivation following repolarization was faster in the mutants that increased and slower in the mutants that decreased the amplitude of the slow component of inactivation. ShB, ●; L396A, ▽; T388A, ◇; L385F, □; L385V, △; L385A, ○.

METHODS. Site-directed mutagenesis, the expression in *Xenopus* oocytes of run-off transcripts of *Shaker* cDNA, two-electrode voltage clamp and patch clamp recording and analysis were done as described previously²⁰. Substitutions are named YxxxZ where Y, the normally occurring amino acid at position xxx of the ShB alternatively spliced variant, has been replaced by amino acid Z. Only oocytes with membrane resistances ≥ 0.5 MΩ which expressed peak currents ≤ 3 μA were included in the two-electrode whole cell recordings. Two-electrode voltage-clamp recordings were done with an MBSH bath solution of (mM) 88 NaCl, 2.4 NaHCO₃, 1 KCl, 0.82 MgSO₄, 0.41 CaCl₂, 0.33 Ca(NO₃)₂, 10 HEPES, pH 7.5 (NaOH), 11 ± 1 °C. Series resistance, approximated at 500 Ω (L. Timpe, unpublished results), was not compensated for, giving errors of up to 1.5 mV with the largest measured currents. Digital subtraction of leak currents was done using scaled averages of currents evoked by ten 20-mV hyperpolarizing pulses from the holding potential of -100 mV. Records were filtered at 1 kHz and sampled at 5 kHz. Recovery from inactivation was determined from two-electrode voltage clamp experiments (11 ± 1 °C) in which a control 100 ms voltage pulse from a holding potential of -100 mV to +60 mV was followed, after an interval of increasing duration at -100 mV, by a second identical depolarization. The peak amplitude minus the amplitude at the end of the current evoked by the second depolarization is expressed as a fraction of the corresponding values of the control pulse and plotted as a function of the inter-pulse recovery interval. Values are mean ± s.e. for three to five oocytes for each mutant.

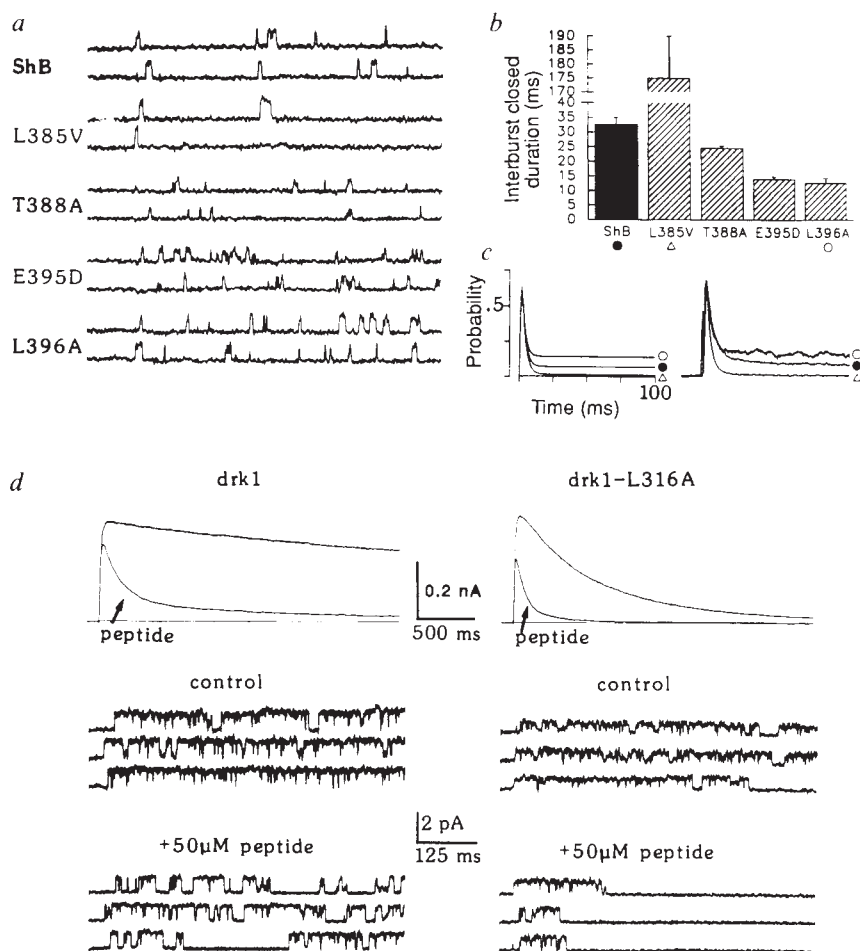


Because cytoplasmic application of quaternary ammonium pore blockers induces fast inactivation in delayed-rectifier channels and interferes with fast inactivation of the A-type *Shaker* channel it has been suggested that fast inactivation may be mediated by a direct occlusion of the permeation pathway by the inactivation gate^{2,17-19}. To examine the possibility that the receptor for the inactivation gate lies near the permeation pathway, we studied the effects of mutations in the S4-S5 loop on single-channel conductance in symmetric K⁺ solutions. Both inward and outward K⁺ currents through the channel were reduced by the subset of mutations that affected inactivation (Table 1), with the most dramatic effect caused by mutations at S392 (Fig. 3c, e, Table 1) and by the alanine substitution of L385 (Table 1). As with L385A in ShB the equivalent mutation L316A in drk1 reduced channel conductance (from 13.0 ± 0.2 pS

to 10.4 ± 0.9 pS; mean ± s.d., N = 4 of each; Fig. 2d). In contrast, there was little effect on channel conductance in the mutants which did not affect channel inactivation (Table 1). The observed reduction of single-channel current amplitude was not due to effects of the mutations on channel gating because only long openings that reached full channel-current amplitude were measured and because the same reduction was observed in double mutants, each with a mutation in the S4-S5 loop combined with the N-terminal deletion mutation ShB(Δ17-25) which greatly increases channel open time²⁰ (Fig. 3b, d; Table 1 legend). The effects on single-channel conductance could not be due to changes of surface charge, as none of the mutations tested altered the charge of the residue. Rather, those residues that are likely to function as a receptor for the inactivation gate probably lie near the permeation pathway, so that even the most

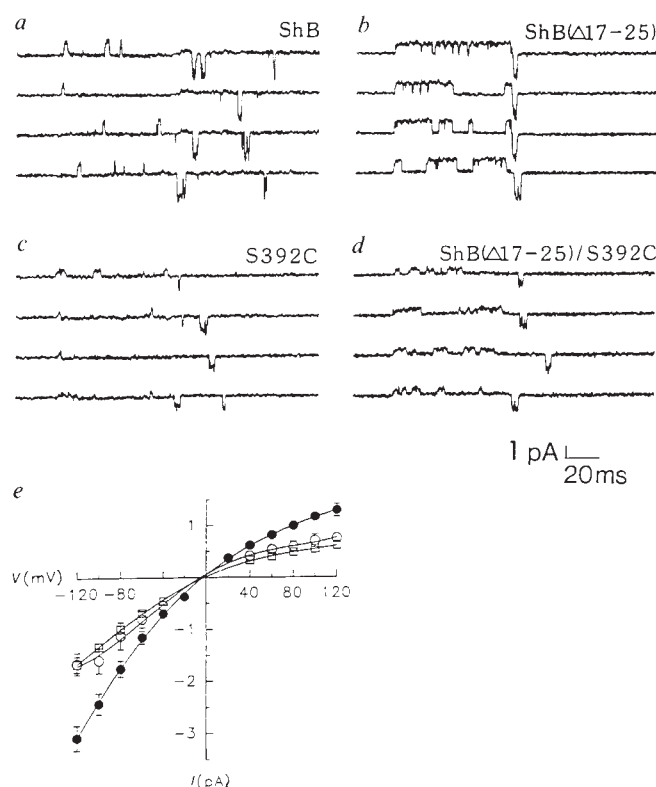
FIG. 2 Mutations in the S4-S5 loop alter the stability of the inactivated state of ShB (a-c) and drk1 (d). a, Pairs of traces evoked by steps from -100 to +100 mV to patches with one active channel of either ShB or one of the four mutants indicated.

b, Changes in the rate of channel reopening owing to mutations in the S4-S5 loop measured from inside-out patches with one functional channel during 160 ms (or 640 ms in the case of L385V) depolarizations to +100 mV. The mean ($\pm$ s.d.) closed duration between bursts of openings was measured as the longer time-constant of two exponential fits to the closed-time distribution. Mean open times and number of openings per burst (expressed as mean $\pm$ s.e. (N)) were ShB: 1.15 ± 0.08 (7) and 1.18 ± 0.04 (4); L385V: 0.90 ± 0.07 (6) and 1.10 ± 0.02 (8); T388A: 0.61 ± 0.05 (7) and 1.57 ± 0.06 (5); and L396A: 1.28 ± 0.20 (3) and 1.46 ± 0.10 (3). c, Changes in the amplitude of the slow component of inactivation can be attributed to increases in the rate of reopenings. Simulations of macroscopic currents at large depolarizations (left) using the model of Zagotta and Aldrich²⁵ with $\alpha = 2,000$, $\beta = 20$, $C \rightarrow O = 3,800$, $O \rightarrow C = 450$, $O \rightarrow I = 400$ and $I \rightarrow O =$ (values presented in b)⁻¹, are similar to normalized macroscopic currents measured from large inside-out patches (right). Symbols indicate identity of mutant as in b. Time base is the same on the left and right. d, Fast inactivation induced in drk1 and the S4-S5 loop mutant drk1-L316A (equivalent to the mutation L385A in ShB) by the cytoplasmic application of 50 μ M of synthetic peptide corresponding to the first 20 amino acids of ShB. Currents were evoked in inside-out patches by steps from -100 mV to +100 mV. Macroscopic currents (top) from inside-out patches before and after bath perfusion of 50 μ M ShB N-terminal peptide. Single-channel records in the attached configuration as control (middle) before excision in the presence of 50 μ M peptide in bath (bottom). Inside-out patch recordings were done in symmetric K⁺ concentration solutions (in mM): internal (bath)



98 KCl, 0.5 MgCl₂, 1 EGTA, 10 HEPES, pH 7.1 (KOH); and external (pipette) 98 KCl, 1.5 MgCl₂, 0.3 CaCl₂, 10 HEPES, pH 7.1 (KOH); (21–24 °C).

FIG. 3 Mutation of S392 in the S4-S5 loop reduces single-channel conductance. a-d, Representative sweeps evoked by biphasic (+100/-100 mV) voltage pulses for wild-type ShB, ShB(Δ 17-25), S392C and the double mutant ShB(Δ 17-25)/S392C. The observation of a reduction in channel conductance in the S392C mutant with a wild-type ShB N terminus (c) was confirmed in the double mutant of S392C/ShB(Δ 17-25) (d). The deletion in the N-terminal region leads to an increased channel open time²⁰ (b, d). e, Open-channel current voltage relations of S392C and S392A show that a parallel reduction occurs in both inward and outward conductance, preserving the inward rectification of wild-type ShB. ShB, ●; S392C, □; S392A, ○. METHODS. Single-channel currents were evoked in inside-out patches by 80 ms depolarizations from a holding potential of -100 mV to positive potentials followed by 80 ms hyperpolarizations to the equivalent negative potential (that is +20/-20 mV, +40/-40 mV and so on). For each patch, mean current amplitude values at each step potential were based on 30-50 manual measurements of openings that were long enough in duration to square off at the top. Plotted values in e represent mean $\pm$ s.d. from six patches for wild-type ShB, five patches each for the mutants. Interpolation of data points was done by fourth order polynomial fits. Inside-out patch recordings were done in symmetric K⁺ solutions as in Fig. 2 (21–24 °C).



conservative alteration of their side chains impedes the K⁺ ion flux through the channel in both directions.

The dual effect on single-channel current amplitude and inactivation of S4-S5 loop mutants distinguishes them from mutations in the N terminus that alter inactivation without affecting single-channel current amplitude (Fig. 3*a, b*)³. Taken together, these observations are consistent with the proposal that the S4-S5 loop forms part of the inactivation receptor and that this receptor lies at or near the cytoplasmic mouth of the permeation pathway for K⁺. Another proposed portion of the permeation pathway, accessible to external^{21,22} and internal^{22,23} tetraethylammonium, is encoded by H5, where amino-acid sub-

stitutions affect channel conductance^{21,22} and ion selectivity²⁴.

The residues in the S4-S5 loop that affect channel conductance and inactivation are distributed in such a way that they would be on the same face of an alpha helix whereas residues without effect would face in other directions. Our study, therefore, suggests that the S4-S5 loop may form an alpha helix with one face lying near the cytoplasmic mouth of the pore where it forms part of the inactivation receptor. If the S4 segment moves outward in response to depolarization as hypothesized⁹⁻¹⁴, then an accompanying conformational change in the attached S4-S5 loop may explain the tight coupling observed between activation and inactivation^{14,25}. □

Received 29 March; accepted 22 July 1991.

- Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, Massachusetts 1984).
- Armstrong, C. M. & Bezanilla, C. F. *J. gen. Physiol.* **70**, 567-590 (1977).
- Hoshi, T., Zagotta, W. N. & Aldrich, R. W. *Science* **250**, 533-538 (1990).
- Zagotta, W. N., Hoshi, T. & Aldrich, R. W. *Science* **250**, 568-571 (1990).
- Schwarz, T. L., Tempel, B. L., Papazian, D. M., Jan, Y. N. & Jan, L. Y. *Nature* **331**, 137-142 (1988).
- Freth, G. C., Van Dongen, A. M. J., Schuster, G., Brown, A. M. & Joho, R. H. *Nature* **340**, 642-645 (1989).
- Zagotta, W. N., Hoshi, T. & Aldrich, R. W. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7243-7243 (1989).
- Isacoff, E. Y., Papazian, D. M., Timpe, L. C., Jan, Y. N. & Jan, L. Y. *Cold Spring Harb. Symp. quant. Biol.* **55**, 9-17 (1990).
- Greenblatt, R. E., Blatt, Y. & Montal, M. *FEBS Lett.* **193**, 125-134 (1985).
- Guy, H. R. & Seetharamulu, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 508-521 (1986).
- Noda, M. *et al. Nature* **320**, 188-192 (1986).
- Catterall, W. A. *Rev. Biochem.* **55**, 953-985 (1986).
- Stuhmer, W. *et al. Nature* **339**, 597-603 (1989).
- Papazian, D. M., Timpe, L. C., Jan, Y. N. & Jan, L. Y. *Nature* **349**, 305-310 (1991).
- Baumann, A., Grupe, A., Ackermann, A. & Pongs, O. *EMBO J.* **7**, 2457-2467 (1988).
- Jan, L. Y. & Jan, Y. N. *Trends Neurosci.* **13**, 415-419 (1990).
- Armstrong, C. M. *J. gen. Physiol.* **54**, 553-575 (1969).
- Armstrong, C. M. *J. gen. Physiol.* **58**, 413-437 (1971).

- Choi, K. L., Aldrich, R. W. & Yellen, G. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5092-5095 (1991).
- Isacoff, E. Y., Jan, Y. N. & Jan, L. Y. *Nature* **345**, 530-534 (1990).
- MacKinnon, R. & Yellen, G. *Science* **250**, 276-279 (1990).
- Hartmann, H. A. *et al. Science* **251**, 942-945 (1991).
- Yellen, G., Jurman, M., Abramson, T. & MacKinnon, R. *Science* **251**, 939-942 (1991).
- Yool, A. & Schwarz, T. L. *Nature* **349**, 700-704 (1991).
- Zagotta, W. N. & Aldrich, R. W. *J. gen. Physiol.* **95**, 29-60 (1990).
- Pfaffinger, P. J., Furukawa, Y., Zhao, B., Dugan, D. & Kandel, E. R. *J. Neurosci.* **11**, 918-927 (1991).
- Ribera, A. B. *Neuron* **5**, 691-701 (1990).
- McCormack, T., Vega-Saenz de Miera, C. & Rudy, B. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5227-5231 (1990).
- Roberds, S. L. & Tamkun, M. M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1798-1802 (1991).
- Guy, H. R. & Conti, F. *Trends Neurosci.* **13**, 201-206 (1990).

ACKNOWLEDGEMENTS. We thank J. Linsman for use of his patch-clamp apparatus, R. Joho for providing the drk1 clone, W. Zagotta and P. Slesinger for providing the simulation program used in Fig. 2*c*, B. Hille for consultation on ion permeation, C. Armstrong, T. Schwarz, P. Slesinger, T. Baldwin, G. Lopez and B. Leckie for discussion, S. Fried for technical assistance, K. Bibbons for participation in the initial phase of this study, C. Turck for peptide synthesis, L. Ackerman for photography, and K. Prewitt for preparation of the manuscript. E.Y.I. is an MDA postdoctoral fellow. Y.N.J. and L.Y.J. are Howard Hughes Investigators. This work was supported by the NIH.

Internal initiation of translation mediated by the 5' leader of a cellular mRNA

Dennis G. Macejak & Peter Sarnow

Molecular Biology Program, Department of Biochemistry, Biophysics and Genetics, and Department of Microbiology and Immunology, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

A RIBOSOME-SCANNING model has been proposed to explain the initiation of eukaryotic messenger RNAs¹ in which binding of the 43S ternary ribosomal subunit near or at the 5' end of the mRNA is facilitated by an interaction between the methylated cap-structure at the end of the mRNA and the cap-binding protein complex eIF-4F (refs 2, 3). But picornaviral mRNAs do not have a 5' terminal cap structure and are translated by internal ribosome binding⁴⁻⁷. A cellular mRNA, encoding the immunoglobulin heavy-chain binding protein, can be translated in poliovirus-infected cells at a time when cap-dependent translation of host cell mRNAs is inhibited⁸. We report here that the 5' leader of the binding protein mRNA can directly confer internal ribosome binding to an mRNA in mammalian cells, indicating that translation initiation by an internal ribosome-binding mechanism is used by eukaryotic mRNAs.

We constructed a eukaryotic expression vector to test the ability of various RNA sequences to mediate internal ribosome binding. The vector pSV_ΔCAT/ICS/LUC should produce dicistronic mRNAs after transfection into eukaryotic cells (Fig. 1*a*), and is similar to the one described by Pelletier and Sonenberg⁴. The first cistron in the dicistronic mRNA, encoding chloramphenicol acetyltransferase (CAT), should be translated by a conventional cap-dependent scanning mechanism. But the second cistron, encoding firefly luciferase (LUC), should be

translated only if preceded by an internal ribosome-binding site, such as the sequences in the 5' leader of poliovirus RNA⁴, in the intercistronic spacer (ICS) region. Three vectors were constructed containing different ICS sequences: sequences 'ED' as a negative control⁹, the 5' noncoding region of poliovirus as a positive control⁴ and the 5' noncoding region of human immunoglobulin heavy-chain binding protein (BiP) mRNA¹⁰. The plasmids were transfected into HeLa cells and, as expected, all three plasmids produced mRNAs from which CAT was translated with comparable efficiency (Fig. 1*b*). Measurement of LUC activities in the same lysates showed that the mRNAs with the 5' noncoding regions of poliovirus or BiP in the ICS could be translated to produce 20-30 times more LUC than the dicistronic CAT/ED/LUC mRNA (Fig. 1*c*).

Various mechanisms could account for the translation of the second cistron in the dicistronic CAT/BiP/LUC mRNA. First, the 5' noncoding region of BiP, but not the ED sequences, may have allowed efficient readthrough of ribosomes past the CAT stop codons and reinitiation of translation of the second LUC cistron. Second, the CAT/BiP/LUC mRNA, but not CAT/ED/LUC mRNA, could have been fragmented in eukaryotic cells, producing functionally monocistronic LUC mRNAs. Alternatively, the 5' noncoding region of BiP could direct internal ribosome binding.

To investigate the possible readthrough of ribosomes from the CAT to the LUC cistron in the CAT/BiP/LUC mRNA, we constructed a vector that could produce dicistronic CAT/BiP/LUC mRNA with a stable RNA hairpin in the 5' noncoding region of CAT (Fig. 2*a*). RNA hairpin structures in 5' noncoding regions have been shown to suppress translation initiation at downstream AUG codons^{11,12}. If LUC translation were due to ribosomal readthrough, suppression of CAT translation should inhibit translation of LUC as well. By contrast, if translation of LUC were initiated by internal ribosome binding, suppression of CAT translation should not affect LUC translation. An RNA hairpin in the 5' noncoding region suppressed translation of CAT but not of LUC (Fig. 2*b*). Thus, the second

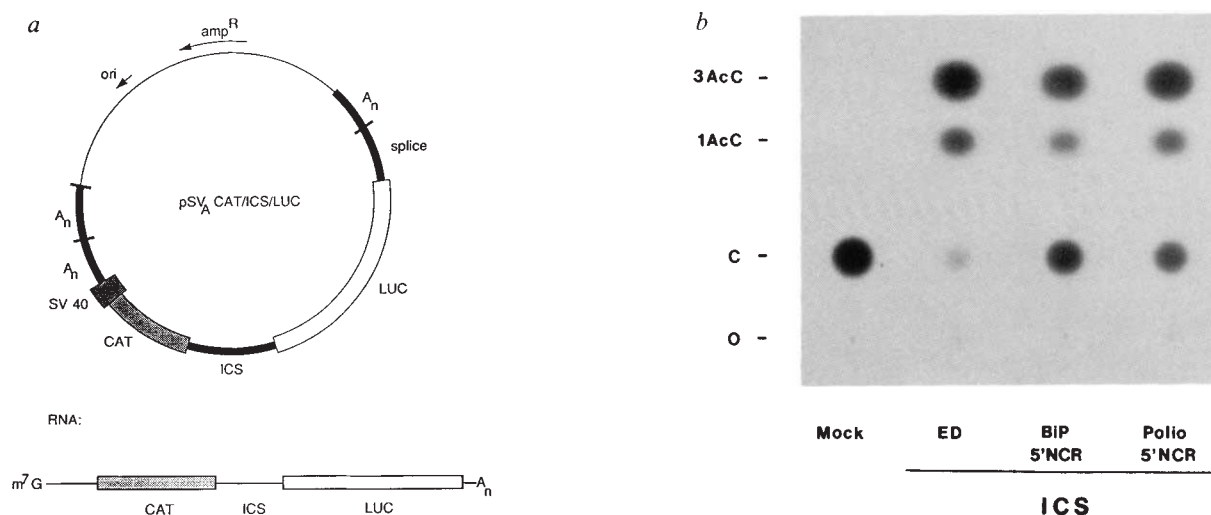


FIG. 1 CAT and LUC expression from dicistronic mRNA molecules. *a*, Diagram of generic vector to express dicistronic mRNAs. The plasmid contains simian virus 40 (SV40) promoter and enhancer, splice and polyadenylation (A_n) signals as indicated. The coding regions of CAT and LUC are separated by an ICS element. *b*, CAT activity obtained from extracts of HeLa cells transfected with dicistronic plasmids containing the following elements in the ICS: a 400-base pair (bp) antisense *Drosophila* Antennapedia complementary DNA (ED)⁹, the 5' noncoding region of BiP cDNA (BiP 5'NCR)¹⁰, or sequences 50–743 of poliovirus type 1 cDNA (polio 5'NCR)⁴. The origin (O), positions of unacetylated CAT (C), and monoacetylated CAT products (1AcC or 3AcC) are indicated. *c*, LUC activity obtained from extracts of HeLa cells transfected with the dicistronic plasmids as described in *b*. LUC activity from two separate experiments is shown after normalizing to CAT activity in the same extract. METHODS. HeLa cells were seeded to 50% confluency 1 day before transfection and 5 μ g plasmid DNA was transfected per 60-mm plate by a calcium phosphate transfection procedure as described¹⁸. Cell extracts were prepared 48 h after transfection and analysed for CAT activity¹⁹ and LUC activity²⁰.

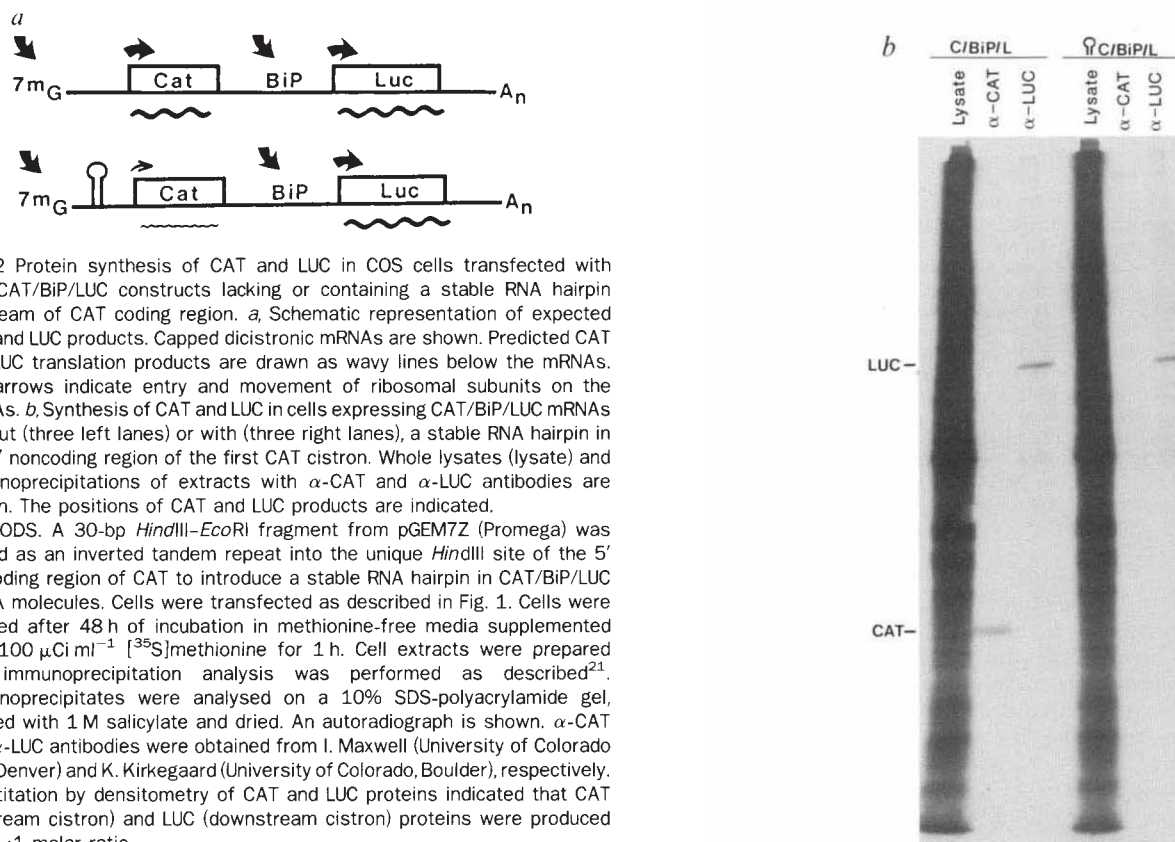


FIG. 2 Protein synthesis of CAT and LUC in COS cells transfected with pSV_ACAT/BiP/LUC constructs lacking or containing a stable RNA hairpin upstream of CAT coding region. *a*, Schematic representation of expected CAT and LUC products. Capped dicistronic mRNAs are shown. Predicted CAT and LUC translation products are drawn as wavy lines below the mRNAs. The arrows indicate entry and movement of ribosomal subunits on the mRNAs. *b*, Synthesis of CAT and LUC in cells expressing CAT/BiP/LUC mRNAs without (three left lanes) or with (three right lanes), a stable RNA hairpin in the 5' noncoding region of the first CAT cistron. Whole lysates (lysate) and immunoprecipitations of extracts with α -CAT and α -LUC antibodies are shown. The positions of CAT and LUC products are indicated. METHODS. A 30-bp *Hind*III-*Eco*RI fragment from pGEM7Z (Promega) was cloned as an inverted tandem repeat into the unique *Hind*III site of the 5' noncoding region of CAT to introduce a stable RNA hairpin in CAT/BiP/LUC mRNA molecules. Cells were transfected as described in Fig. 1. Cells were labelled after 48 h of incubation in methionine-free media supplemented with 100 μ Ci ml⁻¹ [³⁵S]methionine for 1 h. Cell extracts were prepared and immunoprecipitation analysis was performed as described²¹. Immunoprecipitates were analysed on a 10% SDS-polyacrylamide gel, treated with 1 M salicylate and dried. An autoradiograph is shown. α -CAT and α -LUC antibodies were obtained from I. Maxwell (University of Colorado HSC, Denver) and K. Kirkegaard (University of Colorado, Boulder), respectively. Quantitation by densitometry of CAT and LUC proteins indicated that CAT (upstream cistron) and LUC (downstream cistron) proteins were produced at a 1:1 molar ratio.

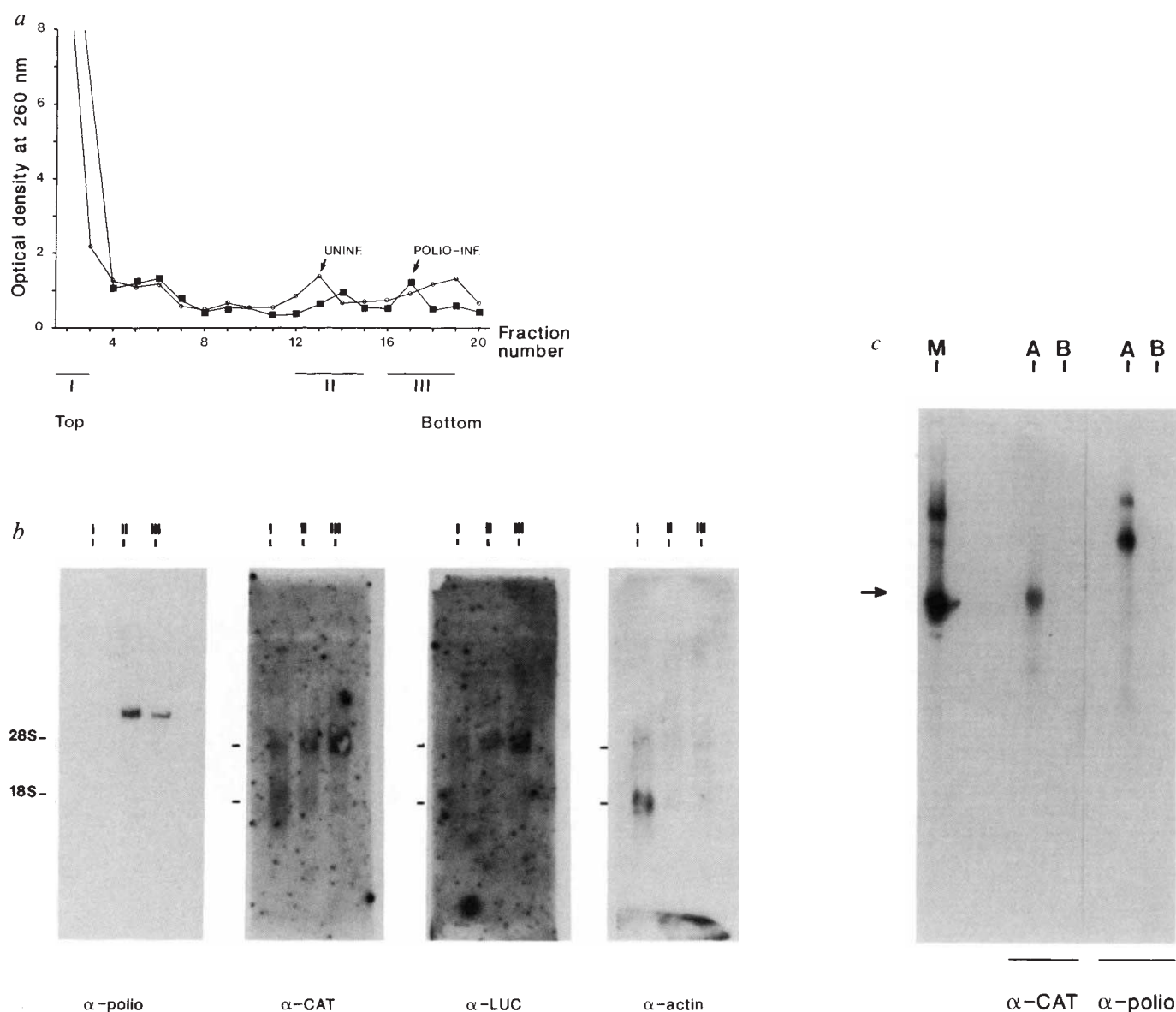


FIG. 3 Polysomal association of CAT/BiP/LUC mRNAs in transfected COS cells. *a*, Sucrose density gradient profiles from extracts of uninfected ($\circ$) or poliovirus-infected ($\blacksquare$) COS cells. COS cells were used because SV40 origin-containing plasmids are amplified in transfected cells allowing a high yield of plasmid-derived transcripts²². Monosomal fractions (I) and polysome-associated fractions (II, III) from infected cell extracts were pooled as indicated. *b*, Analysis of viral and cellular mRNAs in fractions I, II and III by northern blot hybridization. Blots were hybridized with labelled antisense RNAs as indicated. The positions of 18S and 28S ribosomal RNA are denoted at left. The band migrating near 18S rRNA in the non-polysomal fraction I (α -CAT blot) represents crosshybridization of the probe with actin mRNA. *c*, Analysis by northern blot hybridization of viral and cellular mRNAs in polysomal (A) and non-polysomal (B) fractions, obtained after chromatography on Biogel A 1.5 (ref. 15). The arrow indicates the migration of *in vitro* synthesized, dicistronic CAT/BiP/LUC RNA (M). The *in vivo* synthesized dicistronic mRNA runs slightly slower than the *in vitro* synthesized dicistronic mRNA. This is because of the presence of 100 extra

nucleotides (derived from the SV40 promoter) on the 5' leader of CAT.

METHODS. Cells were transfected as for Figs 1 and 2. After 48 h, one set of cells was infected with poliovirus (multiplicity of infection of 50). After an additional 5 h incubation at 37 °C, cytoplasmic RNA was prepared and fractionated by sucrose gradient sedimentation as described²³. Polyadenylate-containing mRNAs from fractions I, II and III were isolated by oligo(dT) affinity chromatography, separated in duplicate on 1% formaldehyde-containing agarose gels and transferred to nitrocellulose. The filters were then hybridized with ³²P-labelled RNAs complementary to CAT or LUC. After autoradiography, the blots were reprobed with ³²P-labelled RNAs complementary to poliovirus mRNA or actin mRNA²⁴. Alternatively, cytoplasmic RNA was prepared and polysomal RNA was separated from nonpolysomal RNA by Biogel A 1.5 chromatography as described¹⁵. Polyadenylate-containing mRNAs from polysomal (A) and nonpolysomal (B) fractions were isolated by oligo(dT) chromatography, separated on agarose gels and processed as described above.

LUC cistron was translated independently of the first CAT cistron, supporting the hypothesis that the 5' noncoding region of BiP can mediate internal ribosome binding.

To address the possible fragmentation of CAT/BiP/LUC mRNAs, we analysed the size and the distribution of the RNA after transfection and subsequent superinfection with poliovirus in mammalian cells. The eIF-4F complex is inactivated in poliovirus-infected cells¹³; as a consequence, capped cellular mRNAs cannot associate with polysomes¹⁴. But translation initiation by internal ribosome binding of the viral mRNA is independent of an intact eIF-4F complex. Therefore, we postu-

lated that intact, dicistronic CAT/BiP/LUC mRNA should be associated with polysomes in poliovirus-infected cells even when eIF-4F-dependent translation of cellular mRNAs is inhibited. Polysomal fractions were prepared by sucrose gradient centrifugation of extracts from cells that had been transfected with pSV₄CAT/BiP/LUC and superinfected with poliovirus (Fig. 3a). As expected, poliovirus mRNA was associated with rapidly sedimenting polysomal fractions (II and III), whereas actin mRNA was predominantly found in the nonpolysomal fraction (I) (Fig. 3b). Radiolabelled RNA probes complementary to both CAT and LUC sequences detected a polyadenylate-containing

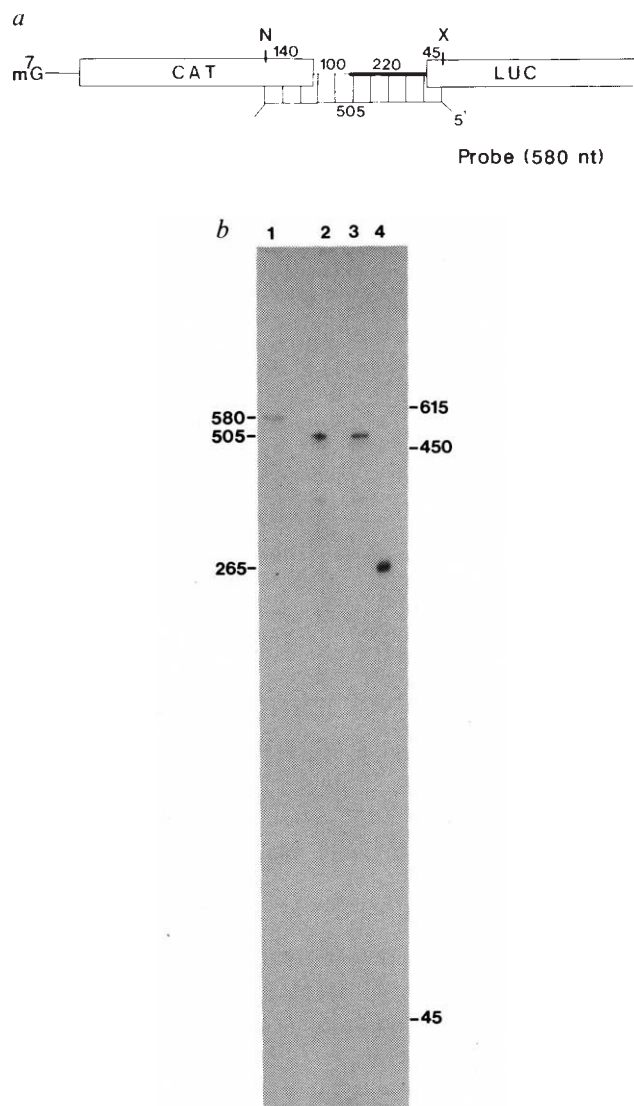


FIG. 4 Ribonuclease-protection analysis of CAT/BiP/LUC mRNA. *a*, Diagram of CAT/BiP/LUC mRNA, hybridized to antisense RNA probe for RNase protection analysis. The 580-nucleotide probe, complementary to the 505 nucleotides of mRNA delimited by the *Nco*I site (N) in CAT and the *Xba*I (X) site in LUC, spans the BiP-containing ICS region. The number of nucleotides protected from each element is indicated. *b*, Ribonuclease protection products. The undigested probe (lane 1), protected fragments from polysome-associated cytoplasmic CAT/BiP/LUC RNA (lane 2), *in vitro* synthesized dicistronic CAT/BiP/LUC RNA (lane 3), and *in vitro* synthesized monocistronic BiP/LUC RNA (lane 4) are shown. The experimentally determined positions of RNA molecules of known lengths (nucleotides) on the same gel are indicated on the right.

METHODS. Polysome-associated cytoplasmic RNA (fraction III, Fig. 3a) was isolated from cells expressing CAT/BiP/LUC RNA containing a stable hairpin upstream of CAT (lane 2). These cells expressed no detectable CAT protein but expressed LUC protein (Fig. 2b). Ribonuclease-protection was performed as described²⁵. Samples were analysed on a 4% polyacrylamide/7M urea gel; an autoradiograph is shown. Both the *in vivo* and *in vitro* dicistronic CAT/BiP/LUC RNAs protected the expected 505-nucleotide fragment (lanes 2 and 3, respectively), in addition to a minor 390-nucleotide RNA. The 390-nucleotide RNA species migrating between the 450- and 265-nucleotide RNA markers (lanes 2 and 3) resulted from ribonuclease A cleavage of the RNA hybrids in an AU-rich region in the CAT-coding region. Hybridization of an oligodeoxynucleotide primer, complementary to the LUC coding region, to polysomal preparations of CAT/BiP/LUC mRNA, followed by primer extension with reverse transcriptase, produced cDNAs extending past the *Nco*I site in the CAT-coding region (not shown), again indicating that the ICS in CAT/BiP/Luc mRNA was intact.

mRNA of the same size. This RNA was present in both polysomal fractions II and III (Fig. 3b). We also purified the polysomes from transfected, superinfected cells by a gel filtration method¹⁵ that separates nonpolysomal RNA and soluble proteins, including ribonucleases, from polysomal RNA. Again, dicistronic mRNA, as well as polioviral mRNA, was associated with the polysomal fraction A (Fig. 3c). These observations indicated that intact, dicistronic CAT/BiP/LUC mRNA was associated with polysomes. In addition, we used a ribonuclease-protection assay (Fig. 4a) to analyse the integrity of the dicistronic CAT/BiP/LUC mRNA further. CAT/BiP/LUC mRNA containing an RNA hairpin in the 5' noncoding region of CAT was isolated from polysomal fraction III after sucrose gradient centrifugation (see Fig. 4b, lane 2) or obtained by *in vitro* synthesis with T7 RNA polymerase (Fig. 4b, lane 3). Dicistronic CAT/BiP/LUC RNAs produced *in vivo*, like the RNA synthesized *in vitro*, contained an intact intercistronic spacer region. Thus, large amounts of functional dicistronic CAT/BiP/LUC RNA were associated with polysomes under conditions in which CAT translation was inhibited (Fig. 2b).

This report provides evidence that the 5' noncoding region of BiP can mediate translation initiation by an internal ribosome-binding mechanism in mammalian cells. Thus, translation by internal ribosome binding is not unique to picornavirus translation. Internal ribosome binding to the poliovirus 5' noncoding region can occur in uninfected as well as in infected cells,

demonstrating that the machinery necessary for this mechanism is present in the host cell. Comparison of the 220-nucleotide 5' leader of BiP mRNA with the 5' leaders of picornaviruses did not reveal any notable sequence identity. This may indicate that undiscovered structural elements in the 5' leader RNA comprise a functional internal ribosome binding site; alternatively, internal ribosome binding sites in viral and cellular mRNAs may be different. Why does the capped cellular BiP mRNA have the potential to be translated by an internal ribosome binding mechanism? It is tempting to speculate that BiP mRNA is bifunctional, and is translated most of the time by a conventional cap-dependent scanning mechanism. But at certain times, an internal ribosome binding mechanism may be advantageous. For example, cap-dependent translation during mitosis in mammalian cells is extremely inefficient, due to the presence of underphosphorylated and therefore nonfunctional eIF-4F (refs 16, 17). An eIF-4F-independent translation initiation mechanism by internal ribosome binding could then be used by mRNAs whose translation is important at this stage of the cell cycle. Thus, even though the picornavirus 5' noncoding regions were the first RNA sequences found to confer internal ribosome binding in mammalian cells⁴⁻⁷, the viruses simply usurped a quite normal host cell mechanism. Thus, internal ribosome binding may also be used by other cellular mRNAs, in addition to BiP mRNA demonstrated here. □

Received 24 May; accepted 15 July 1991.

1. Kozak, M. *Microbiol. Rev.* **47**, 1-45 (1983).
2. Banerjee, A. K. *Microbiol. Rev.* **44**, 175-205 (1980).
3. Sonenberg, N. *Prog. Nucleic Acid Res. molec. Biol.* **35**, 173-297 (1988).
4. Pelletier, J. & Sonenberg, N. *Nature* **334**, 320-325 (1988).
5. Jang, S. K., Davies, M. V., Kaufman, R. J. & Wimmer, E. *J. Virol.* **63**, 1651-1660 (1989).
6. Kaminski, A., Howell, M. T. & Jackson, R. J. *EMBO J.* **9**, 3753-3759 (1990).
7. Belsham, G. J. & Brangwyn, J. K. *J. Virol.* **64**, 5389-5395 (1990).
8. Sarnow, P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5795-5799 (1989).
9. Laughon, A., Boulet, A. M., Bermingham, J. R., Layman, R. A. & Scott, M. P. *Molec. cell. Biol.* **6**, 4676-4689 (1986).
10. Ting, J. & Lee, A. S. *DNA* **7**, 275-286 (1988).
11. Pelletier, J. & Sonenberg, N. *Cell* **40**, 515-526 (1985).
12. Kozak, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2850-2854 (1986).
13. Etchison, D. et al. *J. biol. Chem.* **257**, 14806-14810 (1982).
14. Sonenberg, N. *Adv. Virus Res.* **33**, 175-204 (1987).
15. Calzone, F. J., Angerer, R. C. & Gorovsky, M. A. *Nucleic Acids Res.* **10**, 2145-2161 (1982).
16. Bonneau, A. M. & Sonenberg, N. *J. biol. Chem.* **262**, 11134-11139 (1987).
17. Huang, J. & Schneider, R. J. *Cell* **65**, 271-280 (1991).
18. Ausubel, F. M. et al. in *Current Protocols in Molecular Biology* (eds Ausubel, F. M. et al.) 9.1.1-9.1.7 (Greene and Wiley-Interscience, New York, 1989).
19. Ausubel, F. M. et al. in *Current Protocols in Molecular Biology* (eds Ausubel, F. M. et al.) 9.6.2-9.6.5 (Greene and Wiley-Interscience, New York, 1989).
20. De Wet, J. R. et al. *Molec. cell. Biol.* **7**, 725-737 (1987).
21. Sarnow, P., Ho, Y. S., Williams, J. & Levone, A. J. *Cell* **28**, 387-394 (1982).
22. Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. *Cell* **27**, 279-288 (1981).
23. Katze, M. G., Decorato, D. & Krug, R. M. *J. Virol.* **60**, 1027-1039 (1986).
24. Cleveland, D. W. et al. *Cell* **20**, 95-105 (1980).
25. Kirkegaard, K. & Baltimore, D. *Cell* **47**, 433-443 (1986).

ACKNOWLEDGEMENTS. We thank S. Hambridge, K. Kirkegaard and S.-K. Oh for stimulating discussions and critical reading of the manuscript. This work was supported by the NIH (D.G.M. and P.S.) and in part by the Lucille P. Markey Charitable Trust.

ERRATA

Limited heterogeneity of rearranged T-cell receptor V α transcripts in brains of multiple sclerosis patients

Jorge R. Oksenberg, Simon Stuart, Ann B. Begovich, Robert B. Bell, Henry A. Erlich, Lawrence Steinman & Claude C. A. Bernard

Nature **345**, 344-346 (1990)

AN adenine residue is omitted from position 10 in the legend to Fig. 1 of this paper. The correct sequence of the oligo is:

5' ACGAAGACGAGACCACCGCCCTGC 3'

Other typographical errors were present in Table 1. V α 2 primer should be:

5' GTGTTCCAGAGGGAGCCATTGCC 3'

V α 10 primer should be:

5' ACCCAGCTGCTGGAGCAGAGCCCT 3'

C α primer should be:

5' AATAGGTCGACAGACTTGTCCTGGA 3'

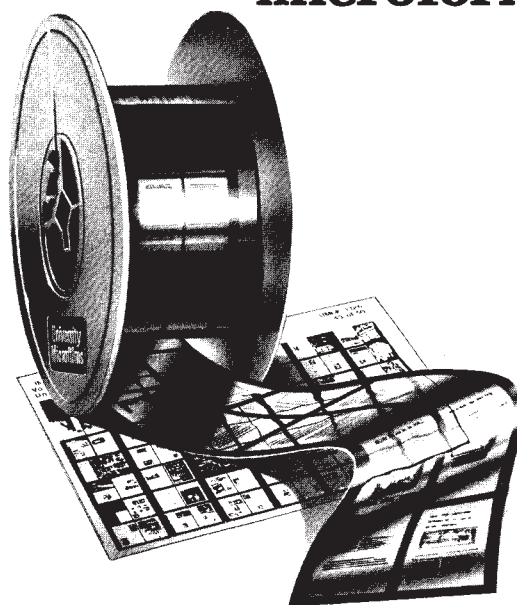
Introduction of a subtle mutation into the *Hox-2.6* locus in embryonic stem cells

Paul Hasty, Ramiro Ramírez-Solis, Robb Krumlauf & Allan Bradley

Nature **350**, 243-246 (1991)

THE Acknowledgements section to this paper should include the statement "The first and second authors had an equal contribution to this work".

nature
is available in
microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone () _____

University
Microfilms
International